Exact Asymptotics for the Random Coding Error Probability

Junya Honda
Graduate School of Frontier Sciences, The University of Tokyo
Kashiwa-shi Chiba 277–8561, Japan
Email: honda@it.k.u-tokyo.ac.jp

***Abstract*—Error probabilities of random codes for memoryless channels are considered in this paper[1]. In the area of communication systems, admissible error probability is very small and it is sometimes more important to discuss the relative gap between the achievable error probability and its bound than to discuss the absolute gap. Scarlett et al. derived a good upper bound of a random coding union bound based on the technique of saddlepoint approximation but it is not proved that the relative gap of their bound converges to zero. This paper derives a new bound on the achievable error probability in this viewpoint for a class of memoryless channels. The derived bound is strictly smaller than that by Scarlett et al. and its relative gap with the random coding error probability (not a union bound) vanishes as the block length increases for a fixed coding rate.**

Keywords*—channel coding, random coding, error exponent, finite-length analysis, asymptotic expansion.

I. Introduction

It is one of the most important task of information theory to clarify the achievable performance of channel codes under finite block length. For this purpose Polyanskiy [2] and Hayashi [3] considered the achievable coding rate under a fixed error probability and a block length. They revealed that the next term to the channel capacity is $\mathrm{O}(1/\sqrt{n})$ for the block length n and expressed by a percentile of a normal distribution.

The essential point for derivation of such a bound is to evaluate error probabilities of channel codes with an accurate form. For this evaluation an asymptotic expansion of sums of random variables is used in [2]. On the other hand, the admissible error probability in communication systems is very small, say, 10^{-10} for example. In such cases it is sometimes more important to consider the *relative* gap between the achievable error probability and its bound than the absolute gap. Nevertheless, an approximation of a tail probability obtained by the asymptotic expansion sometimes results in a large relative gap and it is known that the technique of saddlepoint approximation and the (higher-order) large deviation principle is a more powerful tool rather than the asymptotic expansion [4].

Bounds of the error probability of random codes with a small relative gap have been researched extensively although most of them treat a fixed rate R whereas [2][3] consider varying rate for the fixed error probability. Gallager [5] derived an upper bound called a random coding union bound on the rate of exponential decay of the random coding error probability for fixed rate R. It is proved that this exponent of the random code is tight for both rates below the critical rate [5] and above the critical rate [6].

There have also been many researches on tight bounds of the random coding error probability with vanishing or constant relative error for a fixed rate R. Dobrushin [7] derived a bound of the random coding error probability for symmetric channels in the strong sense that each row and the column of the transition probability matrix are permutations of the others. The relative error of this bound is asymptotically bounded by a constant. In particular, it vanishes in the case that the channel satisfies a nonlattice condition.

For general class of discrete memoryless channels, Gallager [8] derived a bound with a vanishing relative error for the rate below the critical rate based on the technique of exact asymptotics for i.i.d. random variables, and Altuğ and Wagner [9] corrected his result for singular channels. For general (possibly variable) rate R, Scarlett et al. [10] derived a simple upper bound (we write this as $P_{\mathrm{S}}(n)$) of a random coding union bound $P_{\mathrm{RCU}}(n)$ based on the technique of saddlepoint approximation and showed that $P_{\mathrm{RCU}}(n) \le (1+\mathrm{o}(1))P_{\mathrm{S}}(n)$ for nonsingular finite-alphabet discrete memoryless channels [10]. However, This bound does not assure $P_{\mathrm{RCU}}(n) = (1+\mathrm{o}(1))P_{\mathrm{S}}(n)$.

In this paper we consider the error probability P_{RC} of random coding for a fixed but arbitrary rate R below the capacity. We derive a new bound P_{new} which satisfies $P_{\mathrm{new}}(n) = (1+\mathrm{o}(1))P_{\mathrm{RC}}(n)$ for (possibly infinite-alphabet or nondiscrete) nonsingular memoryless channels such that random variables associated with the channels satisfy a condition called a strongly nonlattice condition. The derived bound matches that by Gallager [8] for the rate below the critical rate[2].

The essential point to derive the new bound is that we optimize the parameter depending on the sent and the received sequences $(\boldsymbol{X},\boldsymbol{Y})$ to bound the error probability. This fact contrasts to discussion in [10] and the classic random coding error exponent where the parameter is first fixed and optimized after the expectation over $(\boldsymbol{X},\boldsymbol{Y})$ is taken. We confirm that this difference actually affects the derived bound and by this difference we can assure that the bound also becomes a lower bound of the probability with a vanishing relative error.

[1] This paper is the full version of [1] in ISIT2015 with some corrections and refinements.

[2] In the ISIT proceedings version it was described that the result contradicts the bound in [8] but it was the confirmation error of the author because of the difference of notations between this paper and [11]. See Remark 4 for detail.

II. PRELIMINARY

We consider a memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$. The output distribution for input $x \in \mathcal{X}$ is denoted by $W(\cdot|x)$. Let $X \in \mathcal{X}$ be a random variable with distribution P_X and $Y \in \mathcal{Y}$ be following $W(\cdot|X)$ given X. We define P_Y as the marginal distribution of Y. We assume that $W(\cdot|x)$ is absolutely continuous with respect to P_Y for any x with density

$$\nu(x,y) = \frac{\mathrm{d}W(\cdot|x)}{\mathrm{d}P_Y}(y)\,.$$

We also assume that the mutual information is finite, that is, $I(X;Y) = \mathrm{E}_{XY}[\log \nu(X,Y)] < \infty$.

Let X' be a random variable with the same distribution as X and independent of (X,Y) and define $r(x,y,x') = \log \nu(x',y)/\nu(x,y)$. Since $\nu(X,Y) > 0$ holds almost surely we have $r(X,Y,X') \in \underline{\mathbb{R}} = [-\infty,\infty)$ is well-defined almost surely. $(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{X'}) = ((X_1,\cdots,X_n), (Y_1,\cdots,Y_n), (X_1',\cdots,X_n'))$ denotes n independent copies of (X,Y,X'). We define $r(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{X'}) = \sum_{i=1}^n r(X_i,Y_i,X_i')$.

We consider the error probability of a random code such that each element of codewords $(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_M) \in \mathcal{X}^{n\times M}$ is generated independently from distribution P_X. The coding rate of this code is given by $R = (\log M)/n$. We use the maximum likelihood decoding with ties broken uniformly at random.

A. Error Exponent

Define a random variable $Z(\lambda)$ on the space of functions $\mathbb{R} \to \mathbb{R}$ by

$$Z(\lambda) = \log \mathrm{E}_{X'}\left[\mathrm{e}^{\lambda r(X,Y,X')}\right]$$

and its derivatives by

$$Z^{(m)}(\lambda) = \frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}\log \mathrm{E}_{X'}\left[\mathrm{e}^{\lambda r(X,Y,X')}\right]\,,$$

which we sometimes write by $Z'(\lambda)$, $Z''(\lambda),\cdots$. Here $\mathrm{E}_{X'}$ denotes the expectation over X' for given (X,Y). We define[3]

$$Z(\lambda+\mathrm{i}\xi) = \log \mathrm{E}_{X'}\left[\mathrm{e}^{(\lambda+\mathrm{i}\xi) r(X,Y,X')}\right]$$
$$Z_{\mathrm{a}}(\lambda+\mathrm{i}\xi) = \log \left|\mathrm{E}_{X'}\left[\mathrm{e}^{(\lambda+\mathrm{i}\xi) r(X,Y,X')}\right]\right|\,,$$

where $\lambda,\xi \in \mathbb{R}$ and i is the imaginary unit. Here we always consider the case $\lambda > 0$ and define $\mathrm{e}^{(\lambda+\mathrm{i}\xi)(-\infty)} = 0$. We define

$$Z_i(\lambda) = \log \mathrm{E}_{X'}\left[\mathrm{e}^{\lambda r(X_i,Y_i,X')}\right],\quad \bar{Z}(\lambda) = \frac{1}{n}\sum_{i=1}^n Z_i(\lambda)\,.$$

$Z_{\mathrm{a},i},\, \bar{Z}_{\mathrm{a}}, Z_i^{(m)}$ and $\bar{Z}^{(m)}$ are defined in the same way.

The random coding error exponent for $0 < R < I(X;Y)$ is denoted by

$$\begin{aligned} E_r(R) &= -\inf_{(\alpha,\lambda)\in[0,1]\times[0,\infty)}\{\alpha R + \log \mathrm{E}[\mathrm{e}^{\alpha Z(\lambda)}]\} \\ &= -\min_{\alpha\in(0,1]}\{\alpha R + \log \mathrm{E}[\mathrm{e}^{\alpha Z(1/(1+\alpha))}]\}\,, \end{aligned} \tag{1}$$

and we write the optimal solution of (α,λ) as $(\rho,\eta) = (\rho, 1/(1+\rho))$. We write $\log \mathrm{E}[\mathrm{e}^{\alpha Z(1/(1+\alpha))}] = \Lambda(\alpha)$.

In the strict sense the random coding error exponent represents the supremum of (1) over P_X but for notational simplicity we fix P_X and omit its dependence. See [9, Theorem 2] for a condition that there exists P_X which attains this supremum.

Let P_ρ be the probability measure such that $\mathrm{d}P_\rho/\mathrm{d}P = \mathrm{e}^{\rho Z(\eta)-\Lambda(\rho)}$. We write the expectation under P_ρ by E_ρ and define

$$\begin{aligned} \mu_i &= \mathrm{E}_\rho[Z^{(i)}(\eta)] = \mathrm{e}^{-\Lambda(\rho)}\mathrm{E}[Z^{(i)}(\eta)\mathrm{e}^{\rho Z(\eta)}] \\ \sigma_{ij} &= \mathrm{E}_\rho[(Z^{(i)}(\eta)-\mu_i)(Z^{(j)}(\eta)-\mu_j)] \\ &= \mathrm{e}^{-\Lambda(\rho)}\mathrm{E}[(Z^{(i)}(\eta)-\mu_i)(Z^{(j)}(\eta)-\mu_j)\mathrm{e}^{\rho Z(\eta)}] \\ \Sigma_{ij} &= \begin{pmatrix} \sigma_{ii} & \sigma_{ij} \\ \sigma_{ji} & \sigma_{jj} \end{pmatrix}\,. \end{aligned}$$

From derivatives of $\alpha R + \log \mathrm{E}[\mathrm{e}^{\alpha Z(\lambda)}]$ in α and λ we have

$$\left.\frac{\partial \log \mathrm{E}[\mathrm{e}^{\alpha Z(\eta)}]}{\partial \alpha}\right|_{\alpha=\rho} = \mu_0 \begin{cases} = -R, & \text{if } R \ge R_{\mathrm{crit}}, \\ < -R, & \text{otherwise}, \end{cases} \tag{2}$$

$$\left.\frac{\partial \log \mathrm{E}[\mathrm{e}^{\rho Z(\lambda)}]}{\partial \lambda}\right|_{\lambda=\eta} = \alpha\mu_1 = 0\,. \tag{3}$$

where R_{crit} is the critical rate, that is, the largest R such that the optimal solution of (1) is $\rho = 1$. We assume that $\mu_2 > 0$, or equivalently, $P_Y[|\mathcal{Q}(Y)\setminus\{0\}| > 1] > 0$ where $\mathcal{Q}(y)$ is the support of $\nu(X',y)$. This corresponds to the non-singular assumption in [10][13] for the finite alphabet.

To avoid somewhat technical argument on the continuity and integrability we also assume that there exists $\alpha, b_0 > 0$ and a neighborhood $\mathcal{S}$ of $\lambda = \eta$ such that for any $0 < b_1 < b_2 < 2\pi/h \le \infty$

$$\begin{aligned} &\sup_{\lambda\in\mathcal{S}} \mathrm{E}_\rho[\mathrm{e}^{\alpha|Z^{(m)}(\lambda)|}] < \infty\,,\quad i = 1,2,3, \\ &\sup_{\lambda\in\mathcal{S},\,\xi\in[-b_0,b_0]} \mathrm{E}_\rho[\mathrm{e}^{\alpha|(\partial^4/\partial\xi^4)Z(\lambda+\mathrm{i}\xi)|}] < \infty\,, \\ &\sup_{\lambda\in\mathcal{S},\,\xi\in[b_1,b_2]} \mathrm{E}_\rho[\mathrm{e}^{\alpha|Z_{\mathrm{a}}(\lambda+\mathrm{i}\xi)-Z_{\mathrm{a}}(\lambda)|}] < \infty\,. \end{aligned} \tag{4}$$

where $h \ge 0$ is given later. Note that these conditions trivially hold if the input and output alphabets are finite.

B. Lattice and Nonlattice Distributions

In the asymptotic expansion with an order higher than the central-limit theorem, it is necessary to consider cases that the distribution is lattice or nonlattice separately. Here we call that a random variable $V \in \mathbb{R}^m$ has a lattice distribution if $V \in \{a + \sum_{i=1}^m b_i h_i : \{b_i\} \in \mathbb{Z}^m\}$ almost surely for some $a \in \mathbb{R}^m$ and linearly independent vectors $\{h_i\}_{i=1}^m \in \mathbb{R}^{m\times m}$. For the case $m = 1$ we call the largest h_1 satisfying the above condition the span of the lattice.

On the other hand, we call that $V \in \mathbb{R}^m$ has a strongly nonlattice distribution if $|\mathrm{E}[\mathrm{e}^{\mathrm{i}\langle\xi,V\rangle}]| < 1$ for all $\xi \in \mathbb{R}^m\setminus\{0\}$, where $\langle\cdot,\cdot\rangle$ denotes the inner product. Note that a one dimensional random variable $V \in \mathbb{R}$ is lattice or strongly nonlattice but, in general, there exists a random variable which is not lattice and not strongly nonlattice.

[3] We omit the discussion on the multi-valuedness of $\log z$. The discussion involving logarithm of a complex number in this paper arises by following [12, Sect. XVI.2] and refer this to see that no problem occurs.

As given above, a lattice distribution is defined for a random variable $V \in \mathbb{R}^m$ in standard references such as [14]. In this paper we call that the distribution of $V \in \underline{\mathbb{R}}$ is lattice if the conditional distribution of V given $V > -\infty$ is lattice and nonlattice otherwise. It is easy to see that no contradiction occurs under this definition.

We consider the following condition regarding lattice and nonlattice distributions.

Definition 1. We call that the log-likelihood ratio ν satisfies the lattice condition with span $h > 0$ if the conditional distribution of $\log \nu(X, Y)$ given Y is lattice with span hm_Y almost surely where $m_Y \in \mathbb{N}$ may depend on Y and h is the largest value satisfying this condition.

For notational simplicity we define the span of the lattice for ν to be $h = 0$ if ν does not satisfy the lattice condition. Other than the classification of ν, we also discuss cases that $(Z(\eta), Z'(\eta))$ is strongly nonlattice or not separately.

Note that a one-dimentional random variable $V \in \mathbb{R}$ with support $\mathrm{supp}(V)$ is always lattice if $|\mathrm{supp}(V)| \le 2$, and is strongly nonlattice except for some special cases if $|\mathrm{supp}(V)| \ge 3$. Similarly, a two-dimensional random variable $V \in \mathbb{R}^2$ is always not strongly nonlattice if $|\mathrm{supp}(V)| \le 3$, and is strongly nonlattice except for some special cases if $|\mathrm{supp}(V)| \ge 4$. Based on this observation we see that most channels with input and output alphabet sizes larger than 3 are strongly nonlattice. Another example of each class of channels (excluding those with specially chosen parameters) are given in Table. I.

Remark 1. The above conditions are different from the condition considered in [10] as a classification of lattice and nonlattice cases. This difference arises from two reasons. First, we consider $Z'(\eta)$ in addition to $Z(\eta)$ to derive an accurate bound. Second, the proof of [10, Lemma 1] does not use the correct span when applying the result [15, Sect. VII.1, Thm. 2].

III. Main Result

Define

$$g_h(u) = 1 - \frac{\mathrm{e}^{-\frac{h\eta}{\mathrm{e}^{h\eta}-1}u}(1 - \mathrm{e}^{-h\eta u})}{h\eta u}.$$

for $h \ge 0$. Here we define $(\mathrm{e}^x - 1)/x = (1 - \mathrm{e}^{-x})/x = 1$ for $x = 0$ and therefore $g_0(u) = \lim_{h\downarrow 0} g(u) = 1 - \mathrm{e}^{-u}$. We give some properties on g_h in Appendix A. Now we can represent the random coding error probability as follows.

Theorem 1. *Fix any $0 < R < I(X;Y)$ and $\epsilon > 0$, and let $\delta_2 > 0$ be sufficiently small. Then, for the span $h \ge 0$ of the lattice for ν, there exists $n_0 > 0$ such that for all $n \ge n_0$*

$$\begin{aligned}
&(1-\epsilon)\mathrm{E}\left[g_h\left((1-\epsilon)\frac{\mathrm{e}^{n(\bar{Z}(\eta)+R-(\bar{Z}'(\eta))^2/2(\mu_2-\delta_2))}}{\eta\sqrt{2\pi n\mu_2}}\right)\right]\\
&\le P_{\mathrm{RC}}(n)\\
&\le (1+\epsilon)\mathrm{E}\left[g_h\left((1+\epsilon)\frac{\mathrm{e}^{n(\bar{Z}(\eta)+R-(\bar{Z}'(\eta))^2/2(\mu_2+\delta_2))}}{\eta\sqrt{2\pi n\mu_2}}\right)\right],
\end{aligned}$$

By this theorem we can reduce the evaluation of error probability into that of an expectation over two-dimensional random variable $(\bar{Z}(\eta), \bar{Z}'(\eta))$, although this expectation is still difficult to compute. If $(Z(\eta), Z'(\eta))$ is strongly nonlattice then we can derive the following bound which gives an explicit representation for the asymptotic behavior of P_{RC}.

Theorem 2. *Fix $0 < R < I(X;Y)$ and assume that $(Z(\eta), Z'(\eta))$ has a strongly nonlattice distribution. Then*

$$\begin{aligned}
&P_{\mathrm{RC}}(n)\\
&= \begin{cases}
\frac{\psi_{\rho,h}\mu_2^{(1-\rho)/2}(1+\mathrm{o}(1))}{\eta^\rho(2\pi n)^{(1+\rho)/2}\sqrt{(\mu_2\sigma_{00}+\rho|\Sigma_{01}|)}}\mathrm{e}^{-nE_r(R)}, & R > R_{\mathrm{crit}},\\
\frac{h(1+\mathrm{o}(1))}{2(\mathrm{e}^{\eta h}-1)\sqrt{2\pi n(\mu_2+\sigma_{11})}}\mathrm{e}^{-nE_r(R)}, & R = R_{\mathrm{crit}},\\
\frac{h(1+\mathrm{o}(1))}{(\mathrm{e}^{\eta h}-1)\sqrt{2\pi n(\mu_2+\sigma_{11})}}\mathrm{e}^{-nE_r(R)}, & R < R_{\mathrm{crit}},
\end{cases}
\end{aligned} \tag{5}$$

where

$$\begin{aligned}
\psi_{\rho,h} &= \int_{-\infty}^{\infty} \mathrm{e}^{-\rho w} g_h(\mathrm{e}^w)\mathrm{d}w\\
&= \frac{\Gamma(1-\rho)}{\rho}\left(\frac{h\eta}{\mathrm{e}^{h\eta}-1}\right)^{\rho+1}\frac{\mathrm{e}^h-1}{h}
\end{aligned}$$

for the gamma function Γ.

We prove Theorems 1 and 2 in Sections IV and V, respectively. From this theorem we see that at least for the strongly nonlattice case the error probability of the random coding is

$$P_{\mathrm{RC}}(n) = \begin{cases}\Omega(n^{-(1+\rho)/2}\mathrm{e}^{-nE_r(R)}), & R > R_{\mathrm{crit}}\\ \Omega(n^{-1/2}\mathrm{e}^{-nE_r(R)}), & R \le R_{\mathrm{crit}}.\end{cases} \tag{6}$$

The RHS of (6) for $R > R_{\mathrm{crit}}$ is the same expression as the upper bounds in [10][13] but our bound is tighter in its coefficient and is also assured to be the lower bound.

It may be possible to derive a similar bound as Theorem 2 for the case that $(Z(\eta), Z'(\eta))$ is not strongly nonlattice by replacement of integrals with summations, but for this case the author was not able to find an expression of the asymptotic expansion straightforwardly applicable to our problem and this remains as a future work.

Remark 2. We can show in the same way as Theorem 2 that the random coding *union* bound is obtained by replacement of

TABLE I. Classification of Nonsingular Channels.

		$(Z(\eta), Z'(\eta))$	
		not strongly nonlattice	strongly nonlattice
log-likelihood ratio ν	lattice	BSC	asymmetric BEC
	nonlattice	ternary symmetric channels	binary asymmetric channels

$\psi_{\rho,h}$ with

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\rho w} \min\left\{\frac{h\eta \mathrm{e}^w}{\mathrm{e}^{h\eta}-1}, 1\right\} \mathrm{d}w = \left(\frac{1}{1-\rho}+\frac{1}{\rho}\right)\left(\frac{h\eta}{\mathrm{e}^{h\eta}-1}\right)^{\rho}.$$

On the other hand, the terms $|\rho\Sigma_{01}|$ and σ_{11} in the square roots of (5) are the characteristic parts of the analysis of this paper obtained by the optimization of parameter λ depending on $(\boldsymbol{X}, \boldsymbol{Y})$. Thus, the optimization of λ is necessary to derive a tight coefficient whether we evaluate the error probability itself or the union bound.

Remark 3. The results in this paper assume a *fixed* coding rate R and are weaker in this sense than the result by Scarlett et al. [10] where they assure an upper bound for varying rate by leaving an integral (or a summation) to a form such that the integrant depends on n. It may be possible to extend Theorem 1 for varying rate since the most part of the proof deals with R and the error probability of each codeword separately. However, the proof of Theorem 2 heavily depends on fixed R and it is also an important problem to derive an easily computable bound for varying rate.

Remark 4. In [8] it is shown for discrete nonlattice[4] channels with $R < R_{\mathrm{crit}}$ that

$$P_{\mathrm{RC}}(n) = \frac{(1+\mathrm{o}(1))}{\eta\sqrt{2\pi n \mu_2'}} \mathrm{e}^{-nE_r(R)}, \tag{7}$$

where

$$\mu_2' = \left.\frac{\partial^2 \log \mathrm{E}[\mathrm{e}^{Z(\lambda)}]}{\partial \lambda^2}\right|_{\lambda=\eta} = \frac{2\sum_y (\omega_0(y)\omega_2(y) - \omega_1(y)^2)}{\sum_y \omega_0^2(y)} \tag{8}$$

for

$$\omega_m(y) = \sum_x P_X(x)(\log W(y|x))^m \sqrt{W(y|x)}\,.$$

The author misunderstood that $\mu_2' = \mu_2$ in the ISIT version and described that Theorem 2 contradicts (7). The correct calculation show that $\mu_2' \neq \mu_2$ and

$$\mu_2 = \sigma_{11} = \frac{\sum_y \left(\omega_0(y)\omega_2(y) - \omega_1(y)^2\right)}{\sum_y \omega_0^2(y)}$$

for $(\rho, \eta) = (1, 1/2)$. Therefore no contradiction occurs between this paper and [8].

IV. First Asymptotic Expansion

In this section we give a sketch of the proof of Theorem 1. We prove Theorem 1 separately depending on whether ν satisfies the lattice condition or not. The proofs are different to each other in some places for two reasons. First, we cannot ignore the case that a codeword has the same likelihood as that of the sent codeword under the lattice condition whereas such a case is almost negligible in the nonlattice case. Second, especially in the case of infinite alphabet we have to use the asymptotic expansion with a careful attention to components implicitly assumed to be fixed and the derivation of asymptotic expansion varies in some places between the lattice and nonlattice cases regarding this aspect.

Here we give a proof of Theorem 1 for the case that ν satisfies the lattice condition with span $h > 0$. The proof for the nonlattice case is easier than the lattice case in most places because ties of likelihoods can be almost ignored as described above. See Appendix D for the difference of the proof in the nonlattice case.

Now define

$$\begin{aligned} p_0(\boldsymbol{x}, \boldsymbol{y}) &= P_{\boldsymbol{X}'}[r(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{X}') = 0] \\ p_+(\boldsymbol{x}, \boldsymbol{y}) &= P_{\boldsymbol{X}'}[r(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{X}') > 0] = P_{\boldsymbol{X}'}[r(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{X}') \geq h]\,. \end{aligned} \tag{9}$$

The last equation of (9) holds since $r(x, y, x') = \log \nu(x', y) - \log \nu(x, y)$ and the offset of the lattice of $\log \nu(x', y)$ equals to that of $\log \nu(x, y)$ given y. Under the maximum likelihood decoding, the average error probability P_{RC} is expressed as $P_{\mathrm{RC}} = \mathrm{E}_{\boldsymbol{XY}}[q_M(p_+(\boldsymbol{X}, \boldsymbol{Y}), p_0(\boldsymbol{X}, \boldsymbol{Y}))]$ for

$$\begin{aligned} q_M(p_+, p_0) &= 1 - (1-p_+)^{M-1} \\ &\quad + \sum_{i=1}^{M-1} p_0^i (1 - p_+ - p_0)^{M-i-1} \binom{M-1}{i} \left(1 - \frac{1}{i+1}\right). \end{aligned} \tag{10}$$

Here the first term corresponds to the probability that the likelihood of some codeword exceeds that of the sent codeword, and each component of the second term corresponds to the probability that i codewords have the same likelihood as the sent codeword and the others do not exceed this likelihood.

One of the most basic bound for this quantity is to use a union bound given by

$$q_M(p_+, p_0) \leq \min\{1, (M-1)(p_+ + p_0)\}\,.$$

A lower can also be found in, e.g., [16, Chap. 23]. For evaluation of the error probability with a vanishing relative error the following lemma is useful.

Lemma 1. *It holds for any $c \in (0, 1/2)$ that*

$$\begin{aligned} &\varlimsup_{M\to\infty} \sup_{(p_+,p_0)\in(0,1/3]^2 : p_+ \leq M^c p_0} \frac{q_M(p_+, p_0)}{1 - \frac{\mathrm{e}^{-Mp_+}(1-\mathrm{e}^{-Mp_0})}{Mp_0}} \\ &= \varliminf_{M\to\infty} \inf_{(p_+,p_0)\in(0,1/3]^2 : p_+ \leq M^c p_0} \frac{q_M(p_+, p_0)}{1 - \frac{\mathrm{e}^{-Mp_+}(1-\mathrm{e}^{-Mp_0})}{Mp_0}} = 1\,. \end{aligned}$$

We prove this lemma in Appendix E. We see from this theorem that the error probability can be approximated by

$$1 - \frac{\mathrm{e}^{-Mp_+(\boldsymbol{X},\boldsymbol{Y})}(1 - \mathrm{e}^{-Mp_0(\boldsymbol{X},\boldsymbol{Y})})}{Mp_0(\boldsymbol{X}, \boldsymbol{Y})}$$

for $(\boldsymbol{X}, \boldsymbol{Y})$ satisfying some regularity condition.

Next we consider the evaluation of $p_0(\boldsymbol{X}, \boldsymbol{Y})$ and $p_+(\boldsymbol{X}, \boldsymbol{Y})$. We use Lemma 2 in the following as a fundamental tool of the proof. Let $V_1, \cdots, V_n \in \mathbb{R}$ be (possibly not identically distributed) independent lattice random variables

[4] There is a calculation error for the lattice case in [8] with a redundant factor $\sqrt{\pi}$.

such that the greatest common divisor of their spans[5] is h. Define

$$\Lambda_{V_i}(\lambda) = \log \mathrm{E}[\mathrm{e}^{\lambda V_i}]\,, \quad \Lambda_{\boldsymbol{V}}(\lambda) = \sum_{i=1}^{n} \Lambda_{V_i}(\lambda)\,.$$

Then its large deviation probability is evaluated as follows.

Lemma 2. *Fix $x > \sum_{i=1}^n \mathrm{E}[V_i]$ such that $\Pr[(V_i - x)/h \in \mathbb{Z}] = 1$ and define $\lambda^* > 0$ as the solution of $\Lambda'_{\boldsymbol{V}}(\lambda^*) = x$. Let $\epsilon, \gamma_2, b_0\, \underline{s}_2, \overline{s}_2, \overline{s}_4 > 0$ and $\underline{s}_3, \overline{s}_3 \in \mathbb{R}$ be arbitrary. Then there exists $b_1 = b_1(b_0, \underline{s}_2, \overline{s}_2, \underline{s}_3, \overline{s}_3, \overline{s}_4), n_0 = n_0(\epsilon, b_0, \gamma_2, \underline{s}_2, \overline{s}_2, \underline{s}_3, \overline{s}_3, \overline{s}_4) > 0$ such that*

$$\left| \frac{\Pr[\sum_{i=1}^n V_i = x]}{\frac{h\mathrm{e}^{-n(\eta x - \Lambda_{\boldsymbol{V}}(\lambda^*))}}{\sqrt{2\pi\Lambda''_{\boldsymbol{V}}(\lambda^*)}}} - 1 \right| \le \epsilon\,,$$

$$\left| \frac{\Pr[\sum_{i=1}^n V_i \ge x + h]}{\frac{h\mathrm{e}^{-n(\eta x - \Lambda_{\boldsymbol{V}}(\lambda^*))}}{(\mathrm{e}^{h\lambda^*}-1)\sqrt{2\pi\Lambda''_{\boldsymbol{V}}(\lambda^*)}}} - 1 \right| \le \epsilon\,,$$

hold for all $n \ge n_0$ satisfying

$$n\underline{s}_m \le \sum_{i=1}^{n} \left.\frac{\mathrm{d}^m \Lambda_{V_i}(\lambda)}{\mathrm{d}\lambda^m}\right|_{\lambda=\lambda^*} \le n\overline{s}_m, \qquad i = 2, 3,$$

$$\sum_{i=1}^{n} \left| \frac{\partial^4 \Lambda_{V_i}(\lambda^* + \mathrm{i}\xi)}{\partial \xi^4} \right| \le n\overline{s}_4, \qquad \forall |\xi| \le b_0$$

$$\sum_{i=1}^{n} \left(\log |\mathrm{E}[\mathrm{e}^{(\lambda^*+\mathrm{i}\xi)V_i}]| - \log \mathrm{E}[\mathrm{e}^{\lambda^* V_i}] \right) \le -n\gamma_2, \qquad \forall \xi \in [-\pi/h, \pi/h] \setminus [-b_1, b_1]\,.$$

The proof of this lemma is largely the same as that of [17, Thm. 3.7.4] for the i.i.d. case and given in Appendix B.

Let $b_0, \delta_1, \delta_2, \delta_3, \gamma_1, \gamma_2, \overline{s}_4 > 0$ satisfy $\delta_2 < \min\{\mu_2/2, \mu_2\sqrt{R/12}\}$. To apply Lemma 2 we consider the following sets $\mathcal{A}_m$, $m = 2, 3$, $\mathcal{B}, \mathcal{C}$ to formulate regularity conditions.

$$\begin{aligned}
\mathcal{A}_m &= \{f_1 \in \mathcal{C}_1 : \forall\lambda,\ |f_m(\lambda) - \mu_m| \le \delta_2\}\ ,\\
\mathcal{B} &= \{f_2 \in \mathcal{C}_2 : \forall\lambda, \xi \notin [-b_1, b_1],\ f_2(\lambda, \xi) \le -\gamma_2\}\ ,\\
\mathcal{C} &= \{f_2 \in \mathcal{C}_2 : \forall\lambda, \xi \in [-b_0, b_0],\ f_2(\lambda, \xi) \le \overline{s}_4\}\ ,
\end{aligned}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are the spaces of continuous functions $[\eta - \gamma_1, \eta+\gamma_1] \to \mathbb{R}$ and $[\eta-\gamma_1, \eta+\gamma_1] \times [-\pi/h, \pi/h] \to \mathbb{R}$, respectively, and b_1 is a constant determined from $b_0, \underline{s}_2, \overline{s}_2, \underline{s}_3, \overline{s}_3, \overline{s}_4$ with Lemma 2.

We define the event S as

$$\begin{aligned}
S = {} & \{|\bar{Z}^{(1)}(\eta)| \le \delta_1\} \cup \{\bar{Z}^{(2)}(\lambda) \in \mathcal{A}_2\} \cup \{\bar{Z}^{(3)}(\lambda) \in \mathcal{A}_3\}\\
& \cup \{\bar{Z}_{\mathrm{a}}(\lambda + \mathrm{i}\xi) - \bar{Z}_{\mathrm{a}}(\lambda) \in \mathcal{B}\}\\
& \cup \left\{ \left| \frac{\partial^4}{\partial \xi^4} \bar{Z}^{(4)}(\lambda + \mathrm{i}\xi) \right| \in \mathcal{C} \right\}\ ,
\end{aligned}$$

where we regard $\bar{Z}(\lambda + \mathrm{i}\xi)$ as function $(\lambda, \xi) \mapsto \bar{Z}(\lambda + \mathrm{i}\xi)$. Under this condition we can bound the excess probability of the likelihood of each codeword given the sent codeword $\boldsymbol{X}$ and the received sequence $\boldsymbol{Y}$ as follows.

Lemma 3. *Let $\epsilon > 0$ be arbitrary and $\delta_1 > 0$ in the definition of S be sufficiently small with respect to γ_1. Then, there exists $n_1 > 0$ such that under the event S it holds for all $n \ge n_1$ that,*

$$\begin{aligned}
\frac{h\mathrm{e}^{n(\bar{Z}(\eta) - \bar{Z}'(\eta)^2/2(\mu_2-\delta_2))}}{\sqrt{2\pi n(\mu_2+\delta_2)}}(1-\epsilon) &\le p_0(\boldsymbol{X}, \boldsymbol{Y})\\
&\le \frac{h\mathrm{e}^{n(\bar{Z}(\eta) - \bar{Z}'(\eta)^2/2(\mu_2+\delta_2))}}{\sqrt{2\pi n(\mu_2-\delta_2)}}(1+\epsilon)\,,\\
\frac{h\mathrm{e}^{n(\bar{Z}(\eta) - \bar{Z}'(\eta)^2/2(\mu_2-\delta_2))}}{(\mathrm{e}^{h(\eta+\gamma_1)}-1)\sqrt{2\pi n(\mu_2+\delta_2)}}(1-\epsilon) &\le p_+(\boldsymbol{X}, \boldsymbol{Y})\\
&\le \frac{h\mathrm{e}^{n(\bar{Z}(\eta) - \bar{Z}'(\eta)^2/2(\mu_2+\delta_2))}}{(\mathrm{e}^{h(\eta-\gamma_1)}-1)\sqrt{2\pi n(\mu_2-\delta_2)}}(1+\epsilon)\,.
\end{aligned}$$

Proof: Note that $|\bar{Z}'(\eta)| \le \delta_1$ and $\bar{Z}''(\lambda) \ge \mu_2/2$ for all $\lambda \in [\eta - \gamma_1, \eta + \gamma_1]$ from $\bar{Z}^{(m)}(\lambda) \in \mathcal{A}_m$ and (3). From the convexity of $\bar{Z}(\lambda)$ in λ, if we set $\delta_1 \le \gamma_1\mu_2/2$ then $\bar{Z}(\lambda)$ is minimized at a point in $[\eta - \gamma_1, \eta + \gamma_1]$ with

$$\bar{Z}(\eta) - \frac{(\bar{Z}'(\eta))^2}{2(\mu_2-\delta_2)} \le \min_{\lambda} \bar{Z}(\lambda) \le \bar{Z}(\eta) - \frac{(\bar{Z}'(\eta))^2}{2(\mu_2+\delta_2)}\,.$$

Thus the lemma follows from Lemma 2. ■

Next we define

$$\begin{aligned}
g_h^{(-)}(\boldsymbol{X}, \boldsymbol{Y}) &= (1 - \epsilon/2) g_h\left(\frac{\mathrm{e}^{n(\bar{Z}(\eta)+R-(\bar{Z}'(\eta))^2/2(\mu_2-\delta_2))}}{c^{(-)}\sqrt{n}} \right)\,,\\
g_h^{(+)}(\boldsymbol{X}, \boldsymbol{Y}) &= (1 + \epsilon/2) g_h\left(\frac{\mathrm{e}^{n(\bar{Z}(\eta)+R-(\bar{Z}'(\eta))^2/2(\mu_2+\delta_2))}}{c^{(+)}\sqrt{n}} \right)\,,\\
G_h^{(s)} &= \mathrm{E}[g_h^{(s)}(\boldsymbol{X}, \boldsymbol{Y})],\ s \in \{-, +\}\,,
\end{aligned}$$

where

$$\begin{aligned}
c^{(-)} &= \frac{\eta(\mathrm{e}^{h(\eta+\gamma_1)}-1)\sqrt{2\pi(\mu_2+\delta_2)}}{(\mathrm{e}^{h\eta}-1)(1-\epsilon/2)}\,,\\
c^{(+)} &= \frac{\eta(\mathrm{e}^{h(\eta-\gamma_1)}-1)\sqrt{2\pi(\mu_2-\delta_2)}}{(\mathrm{e}^{h\eta}-1)(1+\epsilon/2)}\,.
\end{aligned}$$

Then the error probability can be evaluated as follows.

Lemma 4. *Fix the coding rate R and assume that the same condition as Lemma 3 holds. Then, for all sufficiently large n,*

$$g_h^{(-)}(\boldsymbol{X}, \boldsymbol{Y}) \le q_M(p_+(\boldsymbol{X}, \boldsymbol{Y}), p_0(\boldsymbol{X}, \boldsymbol{Y})) \le g_h^{(+)}(\boldsymbol{X}, \boldsymbol{Y})\,.$$

This lemma is straightforward from Lemmas 1 and 3. We use the following lemma to evaluate the contribution of the case S^c.

Lemma 5. *Let $\tilde{g}(\boldsymbol{X}, \boldsymbol{Y}) = \mathrm{e}^{n\rho(\bar{Z}(\eta)+R)}$. Then*

$$q_M(p_+(\boldsymbol{X}, \boldsymbol{Y}), p_0(\boldsymbol{X}, \boldsymbol{Y})) \le \tilde{g}(\boldsymbol{X}, \boldsymbol{Y})\,, \tag{11}$$

$$g_h^{(-)}(\boldsymbol{X}, \boldsymbol{Y}) \le \frac{1 + h\eta/2}{(c^{(-)})^{\rho}} \tilde{g}(\boldsymbol{X}, \boldsymbol{Y})\,. \tag{12}$$

[5] The greatest common divisor for a set $\{h_1, h_2, \cdots\}$, $h_i > 0$, is defined as $h > 0$ if h is the maximum number such that $h_i/h \in \mathbb{N}$ for all i and defined as 0 if such h does not exist.

Furthermore, for sufficiently large $\overline{s}_4$ and sufficiently small $\gamma_1 \ll \min\{\delta_2, \delta_3\}$ and $\gamma_2 \ll b_1$ we have

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathrm{E}_{\boldsymbol{XY}}[\mathbb{1}[S^c]\,\tilde{g}(\boldsymbol{X},\boldsymbol{Y})] < -E_r(R)\,.$$

We prove this lemma in Appendix C. The proof is obtained by Cramér's theorem for general topological vector spaces [17, Theorem 6.1.3] with the fact that $\mathcal{C}_1$ and $\mathcal{C}_2$ are separable Banach spaces under the max norm.

Proof of Theorem 1: From Lemma 4, it holds for $\delta_1 \ll \gamma_1 \ll \min\{\delta_2, \delta_3\}$, $\gamma_2 \ll b_1$ and sufficiently large n that

$$\begin{aligned} P_{\mathrm{RC}} &= \mathrm{E}_{\boldsymbol{XY}}[\mathbb{1}[S]\,q_M(p_+(\boldsymbol{X},\boldsymbol{Y}), p_0(\boldsymbol{X},\boldsymbol{Y}))] \\ &\quad + \mathrm{E}_{\boldsymbol{XY}}[\mathbb{1}[S^c]\,q_M(p_+(\boldsymbol{X},\boldsymbol{Y}), p_0(\boldsymbol{X},\boldsymbol{Y}))] \\ &\le G_h^{(+)} + \mathrm{E}_{\boldsymbol{XY}}[\mathbb{1}[S^c]\,q_M(p_+(\boldsymbol{X},\boldsymbol{Y}), p_0(\boldsymbol{X},\boldsymbol{Y}))]\,. \end{aligned}$$

Thus we obtain from Lemma 5 that

$$\frac{P_{\mathrm{RC}}}{G_h^{(+)}} = 1 + \frac{P_{\mathrm{RC}} - G_h^{(+)}}{G_h^{(+)}} \le 1 + \frac{\mathrm{E}_{\boldsymbol{XY}}[\mathbb{1}[S^c]\,\tilde{g}(\boldsymbol{X},\boldsymbol{Y})]}{G_h^{(+)}}\,.$$

Similarly we have

$$\begin{aligned} P_{\mathrm{RC}} &\ge \mathrm{E}_{\boldsymbol{XY}}[\mathbb{1}[S]\,g_h^{(-)}(\boldsymbol{X},\boldsymbol{Y})] \\ &= G_h^{(-)} - \mathrm{E}[\mathbb{1}[S^c]\,g_h^{(-)}(\boldsymbol{X},\boldsymbol{Y})] \end{aligned}$$

and therefore

$$\frac{P_{\mathrm{RC}}}{G_h^{(-)}} \ge 1 - \frac{1+h\eta/2}{(c^{(-)})^\rho}\frac{\tilde{g}(\boldsymbol{X},\boldsymbol{Y})}{G_h^{(-)}}$$

and we see from Lemma 5 and Lemma 6 below that

$$\frac{\tilde{g}(\boldsymbol{X},\boldsymbol{Y})}{G_h^{(s)}} = \mathrm{o}(1),\ s \in \{+,-\}$$

and we obtain Theorem 1. ■

V. Second Asymptotic Expansion

To prove Theorem 2 it is necessary to evaluate the expectation $G_h^{(s)} = \mathrm{E}[g_h^{(s)}(\boldsymbol{X},\boldsymbol{Y})]$. This expectation can be bounded by Lemma 6 below and we give a sketch of its proof in this section.

Lemma 6. *Fix the coding rate $0 < R < I(X;Y)$ assume that $(Z(\eta), Z'(\eta))$ is strongly nonlattice. Then, for any fixed $c_1, c_2 > 0$ and sufficiently large n,*

$$\begin{aligned} &\mathrm{E}\left[g_h\left(\frac{\mathrm{e}^{n(\bar{Z}(\eta)+R-(\bar{Z}'(\eta))^2/2c_1)}}{c_2\sqrt{n}}\right)\right] \\ &= \begin{cases} \frac{\psi_\rho(c_2\sqrt{n})^{-\rho}}{\sqrt{2\pi n(\sigma_{00}+\rho|\Sigma_{01}|/c_1)}}\mathrm{e}^{-nE_r(R)}(1+\mathrm{o}(1)), & R > R_{\mathrm{crit}}, \\ \frac{h\eta(c_2\sqrt{n})^{-1}}{2(\mathrm{e}^{h\eta}-1)\sqrt{1+\sigma_{11}/c_1}}\mathrm{e}^{-nE_r(R)}(1+\mathrm{o}(1)), & R = R_{\mathrm{crit}}, \\ \frac{h\eta(c_2\sqrt{n})^{-1}}{(\mathrm{e}^{h\eta}-1)\sqrt{1+\sigma_{11}/c_1}}\mathrm{e}^{-nE_r(R)}(1+\mathrm{o}(1)), & R < R_{\mathrm{crit}}. \end{cases} \end{aligned}$$

Let Φ_Σ and ϕ_Σ be the cumulative distribution function and the density of a normal distribution with mean zero and covariance Σ, respectively. We define the δ-ball $B_\delta(z) \in \mathbb{R}^2$ around $z \in \mathbb{R}^2$ as $B_\delta(z) = \{z' : \|z-z'\| \le \delta\}$. The oscillation ω_f of f is defined as

$$\begin{aligned} \omega_f(S) &= \sup_{z'\in S} f(z') - \inf_{z'\in S} f(z')\,, \qquad S \subset \mathbb{R}^2\,, \\ \omega_f(\delta;\Phi_\Sigma) &= \sup_{a\in\mathbb{R}^2}\int \omega_f(B_\delta(z))\phi_\Sigma(z+a)\mathrm{d}z\,. \end{aligned}$$

We use the following proposition on the asymptotic expansion for the proof of Lemma 6.

Proposition 1 ([14, Theorem 20.8])**.** *Let $V_1, V_2, \cdots \in \mathbb{R}^2$ be i.i.d. strongly nonlattice random variables with mean zero and covariance matrix Σ. Then, there exists a three-degree polynomial*[6] *$h(z) = h(z_1, z_2)$ such that for any function $f(z)$*

$$\begin{aligned} &\left|\int f(z)\left(1 - \frac{h(z)}{\sqrt{n}}\right)\phi_\Sigma(z)\mathrm{d}z - \mathrm{E}[f(\bar{V})]\right| \\ &\qquad \le \omega_f(\mathbb{R}^2)\delta_n + \omega_f(\delta_n;\Phi_\Sigma)\,, \end{aligned}$$

where δ_n satisfies $\lim_{n\to\infty}\sqrt{n}\delta_n = 0$ and does not depend on f.

To apply this proposition we define

$$f_n(z) = \mathrm{e}^{-\sqrt{n}\rho z_1} g_h\left(\frac{\mathrm{e}^{\sqrt{n}z_1 - z^2/2c_1}}{c_2\sqrt{n}}\right)\,.$$

The oscillations $\omega_{f_n}(\mathbb{R}^2)$ and $\omega_{f_n}(\delta_n;\Phi)$ of f_n are equal to those of

$$\mathrm{e}^{-\sqrt{n}\rho(z_1-\sqrt{n}\Delta)} g_h\left(\frac{\mathrm{e}^{\sqrt{n}(z_1-\sqrt{n}\Delta)-z^2/2c_1)}}{c_2\sqrt{n}}\right)$$

from their definitions.

We can bound the oscillation of f_n as follows.

Lemma 7. *It holds that*

$$\omega_{f_n}(\mathbb{R}^2) = \mathrm{O}(n^{-\rho/2})\,, \tag{13}$$

$$\omega_{f_n}(\delta_n;\Phi) = \mathrm{o}(n^{-\rho/2})\,. \tag{14}$$

Furthermore, if $\rho < 1$ then

$$\omega_f(\delta_n;\Phi) = \mathrm{o}(n^{-(1+\rho)/2})\,. \tag{15}$$

We prove this lemma in Appendix F. By this lemma we can apply Proposition 1 to the proof of Lemma 6, which we give in Appendix G.

VI. Conclusion

We derived a bound of random coding error probability, the relative gap of which converges to zero as the block length increases. The bound applies to any nonsingular memoryless channel such that $(Z(\eta), Z'(\eta))$ is strongly nonlattice. The main difference from other analyses is that we optimize the parameter λ around η depending on the sent and the received sequences $(\boldsymbol{X},\boldsymbol{Y})$. A future work is to extend the bound to the case that $(Z(\eta), Z'(\eta))$ is not strongly nonlattice, that is, $(Z(\eta), Z'(\eta))$ is distributed on a set of lattice points or on a set of parallel lines with an equal interval. It may be possible to derive an expression of asymptotic expansion applicable to our problem by following the discussion in [14, Chap. 5].

[6] The explicit representation of $h(z)$ is given in the original reference [14] but we do not use it in this paper.

Acknowledgment

The author thanks the anonymous reviewers for their helpful comments and suggestion on many related works. This work was supported in part by JSPS KAKENHI Grant Number 26106506.

References

[1] J. Honda, "Exact asymptotics for the random coding error probability," to appear in *ISIT2015*, 2015.

[2] Y. Polyanskiy, H. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[3] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 4947–4966, Nov. 2009.

[4] J. L. Jensen, *Saddlepoint approximations*. Oxford, UK: Oxford University Press, 1995.

[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[6] A. D'yachkov, "Lower bound for ensemble-average error probability for a discrete memoryless channel," *Problems of Information Transmission*, vol. 16, pp. 93–98, 1980.

[7] R. L. Dobrushin, "Asymptotic estimates of the probability of error for transmission of messages over a discrete memoryless communication channel with a symmetric transition probability matrix," *Theory of Probability & Its Applications*, vol. 7, no. 3, pp. 270–300, 1962.

[8] R. G. Gallager, "The random coding bound is tight for the average code." *IEEE Trans. Inform. Theory*, vol. 19, no. 2, pp. 244–246, 1973.

[9] Y. Altuğ and A. Wagner, "Refinement of the random coding bound," *IEEE Trans. Inform. Theory*, vol. 60, no. 10, pp. 6005–6023, Oct 2014.

[10] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "The saddlepoint approximation: Unified random coding asymptotics for fixed and varying rates," in *Proceedings of IEEE International Symposium on Information Theory (ISIT14)*, June 2014, pp. 1892–1896.

[11] R. G. Gallager, *Low-density parity-check codes*, ser. M.I.T. Press research monographs. Cambridge: MIT Press, 1963, no. 21.

[12] W. Feller, *An introduction to probability theory and its applications*, 2nd ed. John Wiley & Sons, 1971, vol. 2.

[13] Y. Altuğ and A. Wagner, "A refinement of the random coding bound," in *Proceedings of 50th Annual Allerton Conference on Communication, Control, and Computing*, Oct 2012, pp. 663–670.

[14] R. Bhattacharya and R. Rao, *Normal Approximation and Asymptotic Expansions*, ser. Classics in Applied Mathematics. SIAM, 1986.

[15] V. Petrov, *Sums of independent random variables*. Springer-Verlag, 1975.

[16] Y. Polyanskiy and Y. Wu, *Lecture notes on Information Theory*, 2012–2014. [Online]. Available: http://people.lids.mit.edu/yp/homepage/data/itlectures_v3.pdf

[17] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 2nd ed., ser. Applications of Mathematics. New York: Springer-Verlag, 1998, vol. 38.

Appendix

A. Properties of Function g_h

Lemma 8. *For $c_h = 1 + h\eta$ it holds that*

$$\begin{aligned} g_h(u) &\le \min\{1, c_h u\} && (16)\\ &\le c_h u^\rho && (17) \end{aligned}$$

and

$$\begin{aligned} 0 \le \frac{\mathrm{d}g_h(u)}{\mathrm{d}u} &\le (u + h\eta)\mathrm{e}^{-u} && (18)\\ &\le c_h\,. && (19) \end{aligned}$$

Proof: We obtain (16) by

$$\begin{aligned} g_h(u) &= 1 - \frac{\mathrm{e}^{-\frac{h\eta}{\mathrm{e}^{h\eta}-1}u}(1-\mathrm{e}^{-h\eta u})}{h\eta u}\\ &\le 1 - \frac{\mathrm{e}^{-u}\mathrm{e}^{-h\eta u}(\mathrm{e}^{h\eta u}-1)}{h\eta u}\\ &\le 1 - \mathrm{e}^{-(1+h\eta)u}\\ &\le \min\{1, c_h u\} \end{aligned}$$

and (17) is straightforward from $0 < \rho \le 1$. We obtain (18) by

$$\begin{aligned} \frac{\mathrm{d}g_h(u)}{\mathrm{d}u} &= \mathrm{e}^{-\frac{h\eta u}{e^{h\eta}-1}}\left(\frac{1-\mathrm{e}^{-h\eta u}}{\mathrm{e}^{h\eta}-1} + \frac{1-\mathrm{e}^{-h\eta u}(1+h\eta u)}{h\eta u^2}\right)\\ &\le \mathrm{e}^{-u}\left(\frac{h\eta u}{h\eta} + \frac{1-(1-h\eta u)(1+h\eta u)}{h\eta u^2}\right)\\ &= (u+h\eta)\mathrm{e}^{-u} \end{aligned}$$

and (19) follows from $u\mathrm{e}^{-u} \le 1$ for any $u \ge 0$. ■

B. Proof of Lemma 2

The proof of Lemma 2 is almost the same as [17, Thm. 3.7.4] where the same result is proved for the i.i.d. case based on the asymptotic expansion for i.i.d. random variables.

In [12, Thm. 2, Sect. XVI], the asymptotic expansion for one-dimensional lattice random variables is derived for i.i.d. cases. It is discussed in [12, Sect. XVI.6.6] that the result is easily extended to non-i.i.d. cases by slightly modifying the proof with some examples depending on regularity conditions. In our setting the following expression is convenient as an asymptotic expansion for non-i.i.d. lattice random variables.

Proposition 2. *Let $\epsilon, \underline{s}_2, \overline{s}_2, \underline{s}_3, \overline{s}_3, \overline{s}_4, b_0, \gamma_2 > 0$ be arbitrary and $V_1, \cdots, V_n \in \mathbb{R}$ be independent lattice random variables such that the greatest common divisor of their spans is h, $\mathrm{E}[V_i] = 0$ and $\Pr[V_i/h \in \mathbb{Z}] = 1$. Then there exists $b_1 = b_1(\underline{s}_2, \overline{s}_2, \underline{s}_3, \overline{s}_3, \overline{s}_4, b_0), n_0 = n_0(\epsilon, \underline{s}_2, \overline{s}_2, \underline{s}_3, \overline{s}_3, \overline{s}_4, b_0, \gamma_2)$ satisfying the following: it holds for all $n \ge n_0$ satisfying*

$$\begin{aligned} &n\underline{s}_2 \le \sum_{i=1}^n V_i^2 \le n\overline{s}_2\,,\\ &n\underline{s}_3 \le \sum_{i=1}^n V_i^3 \le n\overline{s}_3\,,\\ &\sum_{i=1}^n \log|\mathrm{E}[\mathrm{e}^{\mathrm{i}\xi V_i}]| \le -n\gamma_2, \quad \forall \xi \in [-\pi/h, \pi/h] \setminus [-b_1, b_1]\,,\\ &\sum_{i=1}^n \left|\frac{\mathrm{d}^4 \log \mathrm{E}[\mathrm{e}^{\mathrm{i}\xi V_i}]}{\mathrm{d}\xi^4}\right| \le n\overline{s}_4, \qquad \forall|\xi| \le b_0 \end{aligned}$$

that

$$\begin{aligned} \sup_v \Bigg| \Pr\left[\frac{\sum_{i=1}^n V_i}{\sqrt{ns_2}} \le v\right] - \Phi(v) - \frac{s_3}{6\sqrt{n}}(1-v^2)\phi(v)\\ - \phi(v)\tau\left(v, \frac{h}{\sqrt{nA_2}}\right)\Bigg| \le \frac{\epsilon}{\sqrt{n}}\,, \end{aligned}$$

where $s_m = n^{-1}\sum_{i=1}^n V_i^m$, $\tau(v, d) = d\lceil v/d \rceil - v - d/2$, Φ and ϕ are the cumulative distribution function and the density of the standard normal distribution.

Proof of Lemma 2: Let P' be the probability distribution of $\{V_i\}$ such that $\mathrm{d}P'/\mathrm{d}P = \mathrm{e}^{\lambda^* \sum_{i=1}^n V_i}/\mathrm{e}^{\Lambda_{\boldsymbol{V}}(\lambda^*)}$. Then

$$P\left[\sum_{i=1}^n V_i \geq x\right] = \mathrm{e}^{-\Lambda_{\boldsymbol{V}}(\lambda^*)}\mathrm{E}_{P'}\left[\mathrm{e}^{\lambda^* \sum_{i=1}^n V_i}\mathbb{1}\left[\sum_{i=1}^n V_i \geq x\right]\right].$$

Here note that

$$\mathrm{E}_{P'}[V_i] = \frac{\mathrm{E}[V_i \mathrm{e}^{\lambda^* V_i}]}{\mathrm{e}^{\Lambda_{V_i}(\lambda^*)}}$$

and

$$\sum_{i=1}^n \frac{\mathrm{E}[V_i \mathrm{e}^{\lambda^* V_i}]}{\mathrm{e}^{\lambda^* V_i}} = \Lambda'_{\boldsymbol{V}}(\lambda^*) = x$$

from the definition of λ^*. Therefore

$$P\left[\sum_{i=1}^n V_i \geq x\right] = \\ \mathrm{e}^{-\Lambda_{\boldsymbol{V}}(\lambda^*)}\mathrm{E}_{P'}\left[\mathrm{e}^{\lambda^* \sum_{i=1}^n V_i}\mathbb{1}\left[\sum_{i=1}^n (V_i - \mathrm{E}_{P'}[V_i]) \geq 0\right]\right]. \quad (20)$$

Here the variance of V_i under P' are represented by

$$\mathrm{E}_{P'}[(V_i - \mathrm{E}_{P'}[V_i])^2] = \left.\frac{\mathrm{d}^2 \Lambda_{V_i}(\lambda)}{\mathrm{d}\lambda^2}\right|_{\lambda=\lambda^*}$$

and similarly

$$\sum_{i=1}^n \log |\mathrm{E}_{P'}[\mathrm{e}^{\mathrm{i}\xi V_i}]| = \sum_{i=1}^n \log\left|\frac{\mathrm{E}[\mathrm{e}^{\lambda^* V_i}\mathrm{e}^{\mathrm{i}\xi V_i}]}{\mathrm{e}^{\Lambda(\lambda^*)}}\right| \\ = \sum_{i=1}^n \left(\log|\mathrm{E}[\mathrm{e}^{(\lambda^*+\mathrm{i}\xi)V_i}]| - \log \mathrm{E}[\mathrm{e}^{\lambda^* V_i}]\right).$$

Thus we can apply Prop. 2 to the evaluation of (20) and we obtain Lemma 2 by the same argument as [17, Thm. 3.7.4] for the i.i.d. case. ■

C. Proof of Lemma 5

In this appendix we show Lemma 5. Note that (11) is obtained easily by the standard discussion used in the derivation of random coding exponent and (12) also easily follows from (17).

We prove Lemma 5 based on Cramér's theorem in [17] for general vector spaces, which is written for our setting as follows[7].

Proposition 3 (Cramér's theorem [17, Theorem 6.1.3])**.** *Let μ denote the distribution of i.i.d. random variables $V_1, V_2, \cdots$ on a topological real vector space $\mathcal{V}$. Assume that $\mathcal{V}$ is a separable Banach space. Then, for any compact set $\mathcal{S} \subset \mathcal{V}$,*

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \Pr\left[\frac{1}{n}\sum_{i=1}^n V_i \in \mathcal{S}\right] \\ \leq -\inf_{v\in\mathcal{S}} \sup_{\theta\in\mathcal{V}^*}\{\langle v,\theta\rangle - \log \mathrm{E}[\mathrm{e}^{\langle V_1,\theta\rangle}]\},$$

where $\mathcal{V}^*$ is the topological dual of $\mathcal{V}$.

[7] Cramér's theorem in [17] is described for a more general setting such that $\mathcal{V}$ is sufficient to be a metric space under some regularity conditions. When we consider Banach spaces some of these conditions are satisfied and the theorem can be represented in the form of this paper.

We use the following lemma derived from this proposition.

Lemma 9. *Let $\mathcal{V}$ be the space of continuous functions on a compact set $\mathcal{A}$ into $\mathbb{R}$ and $V_1, \cdots, V_n$ be i.i.d. random variables on $\mathcal{V}$ such that $\mathrm{E}[V(s)] = v(s)$ and $\sup_{s\in\mathcal{S}} \mathrm{E}[\mathrm{e}^{\alpha_0|V(s)|}] < \infty$ for some $\alpha_0 > 0$. Then, for any compact set $\mathcal{A}' \subset \mathcal{A}$ and $\epsilon > 0$, the empirical mean $\bar{V} = n^{-1}\sum_{i=1}^n V_i$ satisfies*

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \Pr\left[\sup_{s\in\mathcal{A}'} |\bar{V}(s) - v(s)| \geq \epsilon\right] < 0.$$

Proof: Let $\mathcal{V} \ni f$ be equipped with the max norm

$$\|f\| = \max_{s\in\mathcal{S}} |f(s)|$$

and $\mathcal{V}^*$ be its topological dual, that is, the family of (signed) finite Borel measures on $\mathcal{S}$. Then, we obtain from Cramér's theorem for $\mathcal{S} = \{f \in \mathcal{V} : \sup_{s\in\mathcal{A}'} |f(s) - v(s)| \geq \epsilon\}$ that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \Pr\left[\sup_{s\in\mathcal{A}'} |\bar{V}(s) - v(s)| \geq \epsilon\right] \\ \leq -\inf_{f\in\mathcal{S}} \sup_{\theta\in\mathcal{V}^*}\{\langle f,\theta\rangle - \log \mathrm{E}[\mathrm{e}^{\langle V_1,\theta\rangle}]\}.$$

By considering a set of point mass measures $\{\alpha\delta_{\{s\}} : \alpha \in \mathbb{R}, s \in \mathcal{A}\}$ as a subset of $\mathcal{V}^*$, we obtain

$$\inf_{f\in\mathcal{S}} \sup_{\theta\in\mathcal{V}^*}\{\langle f,\theta\rangle - \log \mathrm{E}[\mathrm{e}^{\langle V_1,\theta\rangle}]\} \\ \geq \inf_{f\in\mathcal{S}} \sup_{s\in\mathcal{A}'} \sup_{\alpha} \left\{\alpha f(s) - \log \mathrm{E}[\mathrm{e}^{\alpha V(s)}]\right\}.$$

Here note that

$$\begin{aligned} 0 &< \frac{\partial^2}{\partial\alpha^2} \log \mathrm{E}\left[\mathrm{e}^{\alpha V(s)}\right] \\ &\leq \frac{\mathrm{E}[V(s)^2 \mathrm{e}^{\alpha V(s)}]}{\mathrm{E}[\mathrm{e}^{\alpha V(s)}]} \\ &\leq \frac{\mathrm{E}[V(s)^2 \mathrm{e}^{\alpha V(s)}]}{\mathrm{E}[1 + \alpha V(s)]} \\ &\leq \frac{\mathrm{E}[V(s)^2 \mathrm{e}^{\alpha|V(s)|}]}{1 - |\alpha|\mathrm{E}[|V(s)|]} \end{aligned}$$

for $|\alpha| < 1/\mathrm{E}[|V(s)|]$. Since there exists $\beta > 0$ such that $x^2\mathrm{e}^{\alpha_0|x|/2} \leq \beta(\mathrm{e}^{\alpha_0|x|} + 1)$ and $|x| \leq \beta\mathrm{e}^{\alpha_0|x|}$ hold for all $x \in \mathbb{R}$,

$$\sup_{|\alpha|<\alpha_0/2} \frac{\partial^2}{\partial\alpha} \log \mathrm{E}\left[\mathrm{e}^{\alpha V(s)}\right] < c$$

for some $c > 0$. Therefore

$$\begin{aligned} &\inf_{f\in\mathcal{S}} \sup_{\theta\in\mathcal{V}^*}\{\langle f,\theta\rangle - \log \mathrm{E}[\mathrm{e}^{\langle V_1,\theta\rangle}]\} \\ &\geq \inf_{f\in\mathcal{S}} \sup_{s\in\mathcal{A}'} \sup_{|\alpha|\leq\alpha_0/2} \left\{\alpha f(s) - \alpha V(s) - c\alpha^2/2\right\} \\ &\geq \inf_{f\in\mathcal{S}} \sup_{|\alpha|\leq\alpha_0} \left\{|\alpha|\epsilon - c\alpha^2/2\right\} \\ &> 0 \end{aligned}$$

and we obtain the lemma. ■

We can apply Lemma 9 to the proof of Lemma 5 from the following lemma.

Lemma 10. *Let $\lambda > 0$ and $\xi \in [-\pi/h, \pi/h]\backslash\{0\}$ be arbitrary. If ν satisfy the lattice condition then*

$$\mathrm{E}_\rho[Z_\mathrm{a}(\lambda + \mathrm{i}\xi)] - \mathrm{E}_\rho[Z_\mathrm{a}(\lambda)] < 0\,.$$

Proof: Let $\mathrm{E}_{X',\lambda}$ be the conditional expectation on X' given (X, Y) under distribution $P_{X',\lambda}$ such that $\mathrm{d}P_{X',\lambda}/\mathrm{d}P_{X'} = \mathrm{e}^{\lambda r(X,Y,X')}/\mathrm{E}_{X'}[\mathrm{e}^{\lambda r(X,Y,X')}]$. Then

$$\begin{aligned}
&\mathrm{E}_\rho[Z(\lambda + \mathrm{i}\xi)] - \mathrm{E}_\rho[Z(\lambda)] \\
&= \mathrm{E}_\rho\left[\log \frac{|\mathrm{E}_{X'}[\mathrm{e}^{(\lambda+\mathrm{i}\xi)r(X,Y,X')}]|}{\mathrm{E}_{X'}[\mathrm{e}^{\lambda r(X,Y,X')}]}\right] \\
&= \mathrm{E}_\rho\left[\log |\mathrm{E}_{X',\lambda}[\mathrm{e}^{\mathrm{i}\xi r(X,Y,X')}]|\right] \\
&= \mathrm{E}_\rho\left[\log |\mathrm{E}_{X',\lambda}[\mathrm{e}^{\mathrm{i}\xi \log \nu(X',Y)}]\mathrm{e}^{-\mathrm{i}\xi \log \nu(X,Y)}|\right] \\
&= \mathrm{E}_\rho\left[\log |\mathrm{E}_{X',\lambda}[\mathrm{e}^{\mathrm{i}\xi \log \nu(X',Y)}]|\right]\,.
\end{aligned}$$

On the other hand, the definition of lattice condition in Def. 1 implies that $P[|\mathrm{E}_{X',\lambda}[\mathrm{e}^{\mathrm{i}\xi \log \nu(X',Y)}]| = 1] < 1$ holds for any $\xi \notin \{2m\pi/h : m \in \mathbb{Z}\}$.

Since P is absolutely continuous with respect to P_ρ we have $P_\rho[|\mathrm{E}_{X',\lambda}[\mathrm{e}^{\mathrm{i}\xi \log \nu(X',Y)}]| = 1] < 1$ for any $\xi \notin \{2m\pi/h : m \in \mathbb{Z}\}$. Thus we obtain $\mathrm{E}_\rho[\log |\mathrm{E}_{X',\lambda}[\mathrm{e}^{\mathrm{i}\xi \log \nu(X',Y)}]|] < 0$ by noting that $\mathrm{E}[V] < 0$ holds for any random variable $V \in \mathbb{R}$ such that $V \le 0$ a.s. and $\Pr[V < 0] > 0$. ∎

Proof of Lemma 5: First we have

$$\begin{aligned}
&\mathrm{E}_{XY}[\mathbb{1}[S^c]\,\mathrm{e}^{n\rho(\bar{Z}(\eta)+R)}] \\
&= \mathrm{e}^{n(\Lambda(\rho)+\rho R)} P_\rho[S^c] \\
&\le \mathrm{e}^{n(\Lambda(\rho)+\rho R)} \Bigg(P_\rho[|\bar{Z}^{(1)}(\eta)| \ge \delta_1] + P_\rho[\bar{Z}^{(2)}(\lambda) \notin \mathcal{A}_2] \\
&\qquad + P_\rho[\bar{Z}^{(3)}(\lambda) \notin \mathcal{A}_3] + P_\rho[\bar{Z}_\mathrm{a}(\lambda + \mathrm{i}\xi) - \bar{Z}_\mathrm{a}(\lambda) \in \mathcal{B}] \\
&\qquad + P_\rho\left[\left|\frac{\partial^4}{\partial \xi^4}\bar{Z}^{(4)}(\lambda + \mathrm{i}\xi)\right| \in \mathcal{C}\right]\Bigg)\,. \qquad (21)
\end{aligned}$$

Note that the moment generating functions of the absolute values of the empirical means in (21) exist from the regularity conditions assumed in (4). It is straightforward from Cramér's inequality that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log P_\rho[|\bar{Z}^{(1)}(\eta)| \ge \delta_1] < 0$$

since $\mathrm{E}_\rho[Z^{(1)}(\eta)] = 0$. It is also straightforward from Lemmas 9 and 10 that the other four probabilities in (21) are exponentially small for sufficiently small γ_1 with respect to (δ_2, δ_3) and

$$\gamma_2 = -\frac{1}{2} \sup_{\substack{\lambda \in [\eta-\gamma_1, \eta+\gamma_1] \\ \xi \in [-\pi/h, \pi/h]\backslash[-b_1, b_1]}} \mathrm{E}_\rho[Z_\mathrm{a}(\lambda + \mathrm{i}\xi) - Z_\mathrm{a}(\lambda)]$$

$$\overline{s}_4 = 2 \sup_{\substack{\lambda \in [\eta-\gamma_1, \eta+\gamma_1] \\ \xi \in [-b_0, b_0]}} \mathrm{E}_\rho\left[\left|\frac{\partial^4 Z(\lambda + \mathrm{i}\xi)}{\partial \xi^4}\right|\right]\,.$$

∎

D. Theorem 1 for Nonlattice Channels

In this appendix we give a brief explanation for the proof of Theorem 1 in the case that $h = 0$, that is, ν does not satisfy the lattice condition. For this case we bound the error probability by

$$\mathrm{E}_{\boldsymbol{XY}}[\tilde{q}_M(\tilde{p}_{1/\sqrt{n}}(\boldsymbol{X}, \boldsymbol{Y}))] \le P_\mathrm{RC} \le \mathrm{E}_{\boldsymbol{XY}}[\tilde{q}_M(\tilde{p}_0(\boldsymbol{X}, \boldsymbol{Y}))]$$

where

$$\begin{aligned}
\tilde{p}_\zeta(\boldsymbol{x}, \boldsymbol{y}) &= P_{\boldsymbol{X}'}[r(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{X}') \ge \zeta] \\
\tilde{q}_M(p) &= 1 - (1-p)^{M-1}\,.
\end{aligned}$$

Similarly to Lemma 1 we have the following lemma.

Lemma 11. *It holds for any $c \in (0, 1/2)$ that*

$$\overline{\lim_{M\to\infty}} \sup_{p \in (0,1/2]} \frac{\tilde{q}_M(p)}{1 - \mathrm{e}^{-pM}} = \underline{\lim}_{M\to\infty} \inf_{p \in (0,1/2]} \frac{\tilde{q}_M(p)}{1 - \mathrm{e}^{-pM}} = 1\,.$$

The proof of this lemma is given in Appendix E. We can obtain Theorem 1 for $h = 0$ by replacing the exact asymptotics for non-i.i.d. lattice random variables with that for nonlattice random variables based on the asymptotic expansion for nonlattice random variables considered in [12, Thm. 1, Sect. XVI]. More precisely we can show Theorem 1 by replacing Prop. 2 with the following proposition, which is also easily obtain from the discussion in [12, Sect. XVI.6.6] for non-i.i.d. random variables.

Proposition 4. *Let $\epsilon, \underline{s}_2, \overline{s}_2, \underline{s}_3, \overline{s}_3, \overline{s}_4, b_0, \gamma_2 > 0$ be arbitrary and $V_1, \cdots, V_n \in \mathbb{R}$ be strongly nonlattice independent random variables such that $\mathrm{E}[V_i] = 0$ and $\Pr[V_i/h \in \mathbb{Z}] = 1$. Then there exists $\underline{d} = \underline{d}(\underline{s}_2, \overline{s}_2, \underline{s}_3, \overline{s}_3, \overline{s}_4, b_0) < \overline{d} = \overline{d}(\epsilon, \underline{s}_2, \overline{s}_2, \underline{s}_3, \overline{s}_3, \overline{s}_4, b_0)$ and $n_0 = n_0(\epsilon, \underline{s}_2, \overline{s}_2, \underline{s}_3, \overline{s}_3, \overline{s}_4, b_0, \gamma_2)$ satisfying the following: it holds for all $n \ge n_0$ satisfying*

$$\begin{aligned}
&n\underline{s}_2 \le \sum_{i=1}^n V_i^2 \le n\overline{s}_2\,, \\
&n\underline{s}_3 \le \sum_{i=1}^n V_i^3 \le n\overline{s}_3\,, \\
&\sum_{i=1}^n \log |\mathrm{E}[\mathrm{e}^{\mathrm{i}\xi V_i}]| \le -n\gamma_2, \quad \forall \xi \in [\underline{d}, \overline{d}]\,, \\
&\sum_{i=1}^n \left|\frac{\mathrm{d}^4 \log \mathrm{E}[\mathrm{e}^{\mathrm{i}\xi V_i}]}{\mathrm{d}\xi^4}\right| \le n\overline{s}_4, \qquad \forall |\xi| \le b_0
\end{aligned}$$

that

$$\begin{aligned}
\sup_v \Bigg| &\Pr\left[\frac{\sum_{i=1}^n V_i}{\sqrt{n s_2}} \le v\right] - \Phi(v) \\
&- \frac{s_3}{6\sqrt{n}}(1 - v^2)\phi(v)\Bigg| \le \frac{\epsilon}{\sqrt{n}}\,.
\end{aligned}$$

E. Bounds on Error Probability for M Codewords

In this appendix we prove Lemmas 1 and 11.

Proof of Lemma 1: First we have

$$\sum_{i=1}^{M-1} p_0^i(1-p_0-p_+)^{M-i-1}\binom{M-1}{i} = (1-p_+)^{M-1}-(1-p_0-p_+)^{M-1} \tag{22}$$

and

$$\begin{aligned}
&\sum_{i=1}^{M-1} p_0^i(1-p_0-p_+)^{M-i-1}\binom{M-1}{i}\frac{1}{i+1}\\
&=\frac{1}{M}\sum_{i=1}^{M-1} p_0^i(1-p_0-p_+)^{M-i-1}\binom{M}{i+1}\\
&=\frac{1}{Mp_0}\sum_{i=2}^{M} p_0^i(1-p_0-p_+)^{M-i}\binom{M}{i}\\
&=\frac{(1-p_+)^M-(1-p_0-p_+)^M}{Mp_0}-(1-p_0-p_+)^{M-1}.
\end{aligned} \tag{23}$$

Combining (22) and (23) with (10) we obtain

$$q_M(p_+,p_0)=1-\frac{(1-p_+)^M-(1-p_0-p_+)^M}{Mp_0}$$

and

$$\begin{aligned}
&1-\frac{q_M(p_+,p_0)}{1-\frac{\mathrm{e}^{-Mp_+}(1-\mathrm{e}^{-Mp_0})}{Mp_0}}\\
&=1-\frac{Mp_0-(1-p_+)^M-(1-p_0-p_+)^M}{Mp_0-\mathrm{e}^{-Mp_+}(1-\mathrm{e}^{-Mp_0})}\\
&=1-\frac{Mp_0-(1-p_+)^M\left(1-\left(1-\frac{p_0}{1-p_+}\right)^M\right)}{Mp_0-\mathrm{e}^{-Mp_+}(1-\mathrm{e}^{-Mp_0})}\\
&=\frac{(1-p_+)^M\left(1-\left(1-\frac{p_0}{1-p_+}\right)^M\right)-\mathrm{e}^{-Mp_+}(1-\mathrm{e}^{-Mp_0})}{Mp_0-\mathrm{e}^{-Mp_+}(1-\mathrm{e}^{-Mp_0})}.
\end{aligned}$$

Here note that $\log(1-x)\ge -x-2x^2$ for $x\le 1/2$. Therefore for $p_0,p_+\le 1/3$ we have

$$\begin{aligned}
&(1-p_+)^M\left(1-\left(1-\frac{p_0}{1-p_+}\right)^M\right)\\
&\le \mathrm{e}^{-Mp_+}\left(1-\mathrm{e}^{-\frac{Mp_0}{1-p_+}-\frac{2Mp_0^2}{(1-p_+)^2}}\right)\\
&\le \mathrm{e}^{-Mp_+}\left(1-\mathrm{e}^{-Mp_0-2Mp_0p_+-5Mp_0^2}\right)\\
&\le \mathrm{e}^{-Mp_+}\left(1-(1-\min\{1,5M(p_+^2+p_+p_0)\})\mathrm{e}^{-Mp_0}\right),
\end{aligned}$$

which implies

$$\begin{aligned}
&\varlimsup_{M\to\infty}\sup_{(p_+,p_0)\in(0,1/3]^2:p_+\le M^c p_0}\left\{1-\frac{q_M(p_+,p_0)}{1-\frac{\mathrm{e}^{-Mp_+}(1-\mathrm{e}^{-Mp_0})}{Mp_0}}\right\}\\
&\le \varlimsup_{M\to\infty}\sup_{p_0\in(0,1/3]}\frac{\min\{1,10M^{1+2c}p_0^2\}}{Mp_0-(1-\mathrm{e}^{-Mp_0})}.\\
&=\varlimsup_{M\to\infty}\sup_{p_0\in(0,1/3]}\frac{1}{M^{1-2c}}\frac{\min\{1,10(Mp_0)^2\}}{Mp_0-(1-\mathrm{e}^{-Mp_0})}\\
&=0.
\end{aligned}$$

Similarly, for $p_0,p_+\le 1/3$ we have

$$\begin{aligned}
&(1-p_+)^M\left(1-\left(1-\frac{p_0}{1-p_+}\right)^M\right)\\
&\ge \mathrm{e}^{-Mp_+-2Mp_+^2}\left(1-\mathrm{e}^{-\frac{Mp_0}{1-p_+}}\right)\\
&\ge \mathrm{e}^{-Mp_+-2Mp_+^2}\left(1-\mathrm{e}^{-Mp_0}\right)\\
&\ge \mathrm{e}^{-Mp_+}(1-\min\{1,2Mp_+^2\})\left(1-\mathrm{e}^{-Mp_0}\right)
\end{aligned}$$

and

$$\begin{aligned}
&\varliminf_{M\to\infty}\inf_{(p_+,p_0)\in(0,1/3]^2:p_+\le M^c p_0}\left\{1-\frac{q_M(p_+,p_0)}{1-\frac{\mathrm{e}^{-Mp_+}(1-\mathrm{e}^{-Mp_0})}{Mp_0}}\right\}\\
&\ge -\varlimsup_{M\to\infty}\sup_{(p_+,p_0)\in(0,1/3]^2:p_+\le M^{1+c}p_0}\frac{\min\{1,2M^{1+2c}p_0^2\}}{Mp_0-(1-\mathrm{e}^{-Mp_0})}\\
&=0,
\end{aligned}$$

which concludes the proof. ■

Proof of Lemma 11: By letting $t(x)=x^{-1}\log(1-x)$ we have

$$\begin{aligned}
\frac{1-(1-p)^{M-1}}{1-\mathrm{e}^{-pM}} &= \frac{1-\mathrm{e}^{p(M-1)t(p)}}{1-\mathrm{e}^{-pM}}\\
&=1-\frac{\mathrm{e}^{-pM}\left(\mathrm{e}^{p(M+(M-1)t(p))}-1\right)}{1-\mathrm{e}^{-pM}}\\
&=1-\frac{\mathrm{e}^{p(M+(M-1)t(p))}-1}{\mathrm{e}^{pM}-1}.
\end{aligned}$$

By $t(x)\le -1$, the second term is bounded from above as

$$\begin{aligned}
\frac{\mathrm{e}^{p(M+(M-1)t(p))}-1}{\mathrm{e}^{pM}-1} &\le \frac{\mathrm{e}^{p}-1}{\mathrm{e}^{pM}-1}\\
&\le \frac{\mathrm{e}^{p}-1}{pM}\\
&\le \frac{\mathrm{e}-1}{M}
\end{aligned} \tag{24}$$

and bounded from below as

$$\begin{aligned}
&\frac{\mathrm{e}^{p(M+(M-1)t(p))}-1}{\mathrm{e}^{pM}-1}\\
&\ge \frac{p(M+(M-1)t(p))}{\mathrm{e}^{pM}-1}\\
&=\frac{M(p+\log(1-p))}{\mathrm{e}^{pM}-1}-\frac{pt(p)}{\mathrm{e}^{pM}-1}\\
&\ge \frac{M(-2p^2)}{\mathrm{e}^{pM}-1} \qquad (25)\\
&\ge -\frac{2}{M}\frac{(Mp)^2}{\mathrm{e}^{pM}-1}\\
&\ge -\frac{2}{M}, \qquad \left(\text{by } \frac{x^2}{\mathrm{e}^x-1}\le 1 \text{ for } x>0\right) \qquad (26)
\end{aligned}$$

where we used $\log(1-p)\ge -p-2p^2$ for $p\in[0,1/2]$ and $t(x)\le 0$ in (25). We complete the proof by letting $M\to\infty$ in (24) and (26). ■

F. Evaluation of Oscillations

In this appendix we prove Lemma 7 on the oscillations of function f_n. We first show Lemmas 12 and 13 below.

Lemma 12. *For any set $S \subset \mathbb{R}^2$,*

$$\omega_{f_n}(S) \le c_h (c_2)^{-\rho} n^{-\rho/2} \sup_{z_2 : z \in S} \mathrm{e}^{-\rho z_2^2/2c_1} .$$

Proof: We can bound f_n as

$$\begin{aligned} f_n(z_1, z_2) &= \mathrm{e}^{-\sqrt{n}\rho z_1} g_h \left(\frac{\mathrm{e}^{\sqrt{n} z_1 - z_2^2/2c_1}}{c_2 \sqrt{n}} \right) \\ &= (c_2\sqrt{n})^{-\rho} \mathrm{e}^{-\rho z_2^2/2c_1} u^{-\rho} g_h(u) \\ &\qquad \left(\text{by letting } u = \tfrac{\mathrm{e}^{\sqrt{n} z_1 - z_2^2/2c_1}}{c_2\sqrt{n}} \right) \\ &\le c_h \left(c_2 \sqrt{n} \right)^{-\rho} \mathrm{e}^{-\rho z_2^2/2c_1} . \quad \text{(by (17))} \end{aligned}$$

Thus we obtain the lemma since $f_n(z) \ge 0$. ■

Lemma 13. *Let $u > 0$ and $r \in [-1/2, 1/2]$ be arbitrary. Then*

$$|g_h((1+r)u) - g_h(u)| \le c_h |r| u \,, \tag{27}$$
$$|g_h((1+r)u) - g_h(u)| \le c_h |r| \,. \tag{28}$$

Proof: Eq. (27) is straightforward from (19). We obtain (28) from

$$\begin{aligned} \frac{\mathrm{d} g_h((1+r)u)}{\mathrm{d}r} &= u \frac{\mathrm{d} g_h(v)}{\mathrm{d}v} \bigg|_{v=(1+r)u} \\ &\le u((1+r)u + h\eta)\mathrm{e}^{-u} \quad \text{(by (18))} \\ &\le 6\mathrm{e}^{-2} + h\eta \mathrm{e}^{-1} \\ &\le c_h . \end{aligned}$$

■

By using these lemmas we can evaluate the oscillation of f_n within a ball as follows.

Lemma 14. *Assume $|z_2| \le c_1\sqrt{n}/2$. Then, for sufficiently large n,*

$$\omega_{f_n}(B_{\delta_n}(z)) \le \frac{8c_h}{c_2} \delta_n \mathrm{e}^{(1-\rho)\sqrt{n} z_1} , \tag{29}$$
$$\omega_{f_n}(B_{\delta_n}(z)) \le 4(1+c_h)\sqrt{n}\delta_n \mathrm{e}^{-\rho\sqrt{n} z_1} . \tag{30}$$

Proof: First we obtain for z' satisfying $\|z' - z\| \le \delta_n$ and sufficiently large n that

$$\begin{aligned} |(z_2')^2 - z_2^2| &\le |z_2' - z_2|(|z_2'| + |z_2|) \\ &\le |z_2' - z_2|(2|z_2| + |z_2 - z_2'|) \\ &\le \delta_n \left| c_1\sqrt{n} + \delta_n \right| \\ &\le 2c_1 \delta_n \sqrt{n} . \quad (\text{by } \lim_{n\to\infty} \delta_n = 0) \end{aligned}$$

Let $w = z_1 - z_2^2/(2c_1\sqrt{n})$ and $w' = z_1' - (z_2')^2/(2c_1\sqrt{n})$. Then

$$\begin{aligned} |w' - w| &\le |z_1' - z_1| + \frac{|z_2^2 - (z_2')^2|}{2c_1\sqrt{n}} \\ &\le 2\delta_n . \end{aligned}$$

Therefore we obtain for sufficiently large n that

$$\left| \frac{\mathrm{e}^{\rho\sqrt{n} w'}}{\mathrm{e}^{\rho\sqrt{n} w}} - 1 \right| \le 2(\rho\sqrt{n}|w' - w|) \le 2\sqrt{n}\delta_n$$
$$\left| \frac{\mathrm{e}^{\rho\sqrt{n} z_1'}}{\mathrm{e}^{\rho\sqrt{n} z_1}} - 1 \right| \le 2\sqrt{n}\delta_n$$

since $\lim_{n\to\infty} \sqrt{n}\delta_n = 0$. Therefore by letting $\delta_n' = 2\delta_n\sqrt{n}$ and using (27) we obtain for sufficiently large n that

$$\begin{aligned} f_n(z') &\le (1+\delta_n')\mathrm{e}^{-\rho\sqrt{n} z_1} g_h \left(\frac{(1+\delta_n')\mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \right) \\ &\le (1+\delta_n')\mathrm{e}^{-\rho\sqrt{n} z_1} \left(g_h \left(\frac{\mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \right) + \frac{c_h \delta_n' \mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \right) , \\ f_n(z') &\ge (1-\delta_n')\mathrm{e}^{-\rho\sqrt{n} z_1} \left(g_h \left(\frac{\mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \right) - \frac{c_h \delta_n' \mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \right) . \end{aligned}$$

We obtain (29) from these inequalities by

$$\begin{aligned} \omega_{f_n}(B_{\delta_n}(z)) &\le 2\delta_n' \mathrm{e}^{-\rho\sqrt{n} z_1} \left(g_h \left(\frac{\mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \right) + \frac{c_h \mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \right) \\ &\le 4\delta_n' \mathrm{e}^{-\rho\sqrt{n} z_1} \frac{c_h \mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \quad \text{(by (16))} \\ &\le 4\delta_n' \mathrm{e}^{(1-\rho)\sqrt{n} z_1} \frac{c_h}{c_2\sqrt{n}} . \end{aligned}$$

Similarly we obtain from (28) that

$$\begin{aligned} f_n(z') &\le (1+\delta_n')\mathrm{e}^{-\rho\sqrt{n} z_1} \left(g_h \left(\frac{\mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \right) + c_h \delta_n' \right) \\ f_n(z') &\ge (1-\delta_n')\mathrm{e}^{-\rho\sqrt{n} z_1} \left(g_h \left(\frac{\mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \right) - c_h \delta_n' \right) . \end{aligned}$$

From these inequalities we obtain (30) by

$$\begin{aligned} \omega_{f_n}(B_{\delta_n}(z)) &\le 2\delta_n' \mathrm{e}^{-\rho\sqrt{n} z_1} \left(g_h \left(\frac{\mathrm{e}^{\sqrt{n} w}}{c_2\sqrt{n}} \right) + c_h \right) \\ &\le 2(1+c_h)\delta_n' \mathrm{e}^{-\rho\sqrt{n} z_1} . \end{aligned}$$

■

Proof of Lemma 7: Let b_n be such that

$$\mathrm{e}^{\sqrt{n} b_n} = n^{1/2 + 1/4\rho} \delta_n^{1/2\rho} .$$

First we have

$$\begin{aligned} &\int \omega_f(\{z' : \|z' - z\| \le \delta\}) \phi_\Sigma(z+a) \mathrm{d}z \\ &\le \int_{|z_2| \le c_1\sqrt{n}/2,\, z_1 \le b_n} \omega_f(B_{\delta_n}(z)) \phi_\Sigma(z+a) \mathrm{d}z \\ &\quad + \int_{|z_2| \le c_1\sqrt{n}/2,\, z_1 \ge b_n} \omega_f(B_{\delta_n}(z)) \phi_\Sigma(z+a) \mathrm{d}z \\ &\quad + \int_{|z_2| \ge c_1\sqrt{n}/2} \omega_f(B_{\delta_n}(z)) \phi_\Sigma(z+a) \mathrm{d}z \\ &\le \int_{|z_2| \le c_1\sqrt{n}/2,\, z_1 \le b_n} \frac{4c_4 c_h}{c_2} \delta_n \mathrm{e}^{(1-\rho)\sqrt{n} z_1} \phi_\Sigma(z+a) \mathrm{d}z \end{aligned}$$

$$
\begin{aligned}
&+\int_{|z_2|\le c_1\sqrt{n}/2, z_1\ge b_n} 2c_4(1+c_h)\sqrt{n}\delta_n \mathrm{e}^{-\rho\sqrt{n}z_1}\phi_\Sigma(z+a)\mathrm{d}z\\
&+\int_{|z_2|\ge c_1\sqrt{n}/2} c_h(c_2\sqrt{n})^{-\rho}\mathrm{e}^{-c_1\rho n/8}\phi_\Sigma(z+a)\mathrm{d}z\\
&\qquad\qquad\text{(by Lemmas 12 and 14)}\\
&\le \int_{-\infty}^{b_n}\frac{4c_4c_h}{c_2}\delta_n\mathrm{e}^{(1-\rho)\sqrt{n}z_1}\phi_{\sigma_{11}}(z_1+a_1)\mathrm{d}z_1\\
&+\int_{b_n}^{\infty}2c_4(1+c_h)\sqrt{n}\delta_n\mathrm{e}^{-\rho\sqrt{n}z_1}\phi_{\sigma_{11}}(z_1+a_1)\mathrm{d}z_1\\
&+\mathrm{o}(n^{-(1+\rho)/2}). \qquad (31)
\end{aligned}
$$

Here recall that $\lim_{n\to\infty}\sqrt{n}\delta_n=0$ and therefore the second term of (31) is bounded as

$$
\begin{aligned}
&\int_{b_n}^{\infty}\sqrt{n}\delta_n\mathrm{e}^{-\rho\sqrt{n}z_1}\phi_{\sigma_{11}}(z_1+a_1)\mathrm{d}z_1\\
&\le\frac{1}{\sqrt{2\pi\sigma_1^2}}\int_{b_n}^{\infty}\sqrt{n}\delta_n\mathrm{e}^{-\rho\sqrt{n}z_1}\mathrm{d}z_1\\
&=\frac{1}{\sqrt{2\pi\sigma_1^2}}\frac{\delta_n\mathrm{e}^{-\rho\sqrt{n}b_n}}{\rho}\\
&=\frac{1}{\sqrt{2\pi\sigma_1^2}}\frac{\delta_n(n^{1/2+1/4\rho}\delta_n^{1/2\rho})^{-\rho}}{\rho}\\
&=\frac{1}{\sqrt{2\pi\sigma_1^2}}\frac{(\sqrt{n}\delta_n)^{1/2}}{\rho n^{(1+\rho)/2}}\\
&=\mathrm{o}(n^{-(1+\rho)/2}).
\end{aligned}
$$

We obtain (14) since the first term of (31) is bounded as

$$
\begin{aligned}
&\int_{-\infty}^{b_n}\delta_n\mathrm{e}^{(1-\rho)\sqrt{n}z_1}\phi_{\sigma_{11}}(z_1+a_1)\mathrm{d}z_1\\
&\le\delta_n\mathrm{e}^{(1-\rho)\sqrt{n}b_n}\\
&=\delta_n(\sqrt{n}(\sqrt{n}\delta_n)^{1/2\rho})^{(1-\rho)}\\
&=n^{-\rho}(\sqrt{n}\delta_n)(\sqrt{n}\delta_n)^{(1-\rho)/2\rho}\\
&=\mathrm{o}(n^{-\rho}).
\end{aligned}
$$

We obtain (15) since the first term of (31) is also bounded for $\rho<1$ as

$$
\begin{aligned}
&\int_{-\infty}^{b_n}\delta_n\mathrm{e}^{(1-\rho)\sqrt{n}z_1}\phi_{\sigma_{11}}(z_1+a_1)\mathrm{d}z_1\\
&\le\frac{1}{\sqrt{2\pi\sigma_{11}^2}}\int_{-\infty}^{b_n}\delta_n\mathrm{e}^{(1-\rho)\sqrt{n}z_1}\mathrm{d}z_1\\
&=\frac{1}{\sqrt{2\pi\sigma_{11}^2}}\frac{\delta_n(\sqrt{n}(\sqrt{n}\delta_n)^{1/2\rho})^{1-\rho}}{(1-\rho)\sqrt{n}}\\
&=\frac{1}{\sqrt{2\pi\sigma_{11}^2}}\frac{n^{-(1+\rho)/2}(\sqrt{n}\delta_n)(\sqrt{n}\delta_n)^{(1-\rho)/2\rho}}{1-\rho}\\
&=\mathrm{o}(n^{-(1+\rho)/2}).
\end{aligned}
$$

■

G. Proof of Lemma 6

First we have

$$
\begin{aligned}
&\mathrm{E}\left[g_h\left(\frac{\mathrm{e}^{n(\bar{Z}(\eta)+R-(\bar{Z}'(\eta))^2/2c_1)}}{c_2\sqrt{n}}\right)\right]\\
&=\mathrm{e}^{n\Lambda(\rho)}\mathrm{E}_\rho\left[\mathrm{e}^{-n\rho\bar{Z}(\eta)}g_h\left(\frac{\mathrm{e}^{n(\bar{Z}(\eta)+R-(\bar{Z}'(\eta))^2/2c_1)}}{c_2\sqrt{n}}\right)\right].
\end{aligned}
$$

Here recall that $\mathrm{E}_\rho[\bar{Z}(\eta)]=\mu_0\le -R$ and $\mathrm{E}_\rho[\bar{Z}'(\eta)]=\mu_1=0$ from (2). By letting $\Delta=-(R+\mu_0)$, we have $\Delta=0$ for $R\ge R_{\mathrm{crit}}$ and $\Delta>0$ for $R<R_{\mathrm{crit}}$. Normalizing $\bar{Z}(\eta)$ and $\bar{Z}'(\eta)$ as $\tilde{Z}_1=\sqrt{n}(\bar{Z}(\eta)+R+\Delta)$ and $\tilde{Z}_2=\sqrt{n}\bar{Z}'(\eta)$, respectively, we have

$$
\begin{aligned}
&\mathrm{E}\left[g_h\left(\frac{\mathrm{e}^{n(\bar{Z}(\eta)+R-(\bar{Z}'(\eta))^2/2c_1)}}{c_2\sqrt{n}}\right)\right]\\
&=\mathrm{e}^{-nE_r(R)}\\
&\quad\cdot\mathrm{E}_\rho\left[\mathrm{e}^{-\sqrt{n}\rho(\tilde{Z}_1-\sqrt{n}\Delta)}g_h\left(\frac{\mathrm{e}^{\sqrt{n}(\tilde{Z}_1-\sqrt{n}\Delta)-\tilde{Z}^2/2c_1)}}{c_2\sqrt{n}}\right)\right].
\end{aligned}
$$

We obtain from Prop. 1 that

$$
\begin{aligned}
&\mathrm{E}_\rho\left[\mathrm{e}^{-\sqrt{n}\rho(\tilde{Z}_1-\sqrt{n}\Delta)}g\left(\frac{\mathrm{e}^{\sqrt{n}(\tilde{Z}_1-\sqrt{n}\Delta)-\tilde{Z}^2/2c_1)}}{c_2\sqrt{n}}\right)\right]\\
&=\iint\frac{\mathrm{e}^{-z^T\Sigma_{01}^{-1}z/2}}{2\pi\sqrt{|\Sigma|}}\left(1-\frac{h(z)}{\sqrt{n}}\right)\mathrm{e}^{-\sqrt{n}\rho(z_1-\sqrt{n}\Delta)}\\
&\quad\cdot g\left(\frac{\mathrm{e}^{\sqrt{n}(z_1-\sqrt{n}\Delta)-z_2^2/2c_1)}}{c_1\sqrt{n}}\right)\mathrm{d}z_1\mathrm{d}z_2+\omega_{f_n}(\delta_n;\Phi).
\end{aligned}
$$

For the case (i) $\rho<1,\Delta=0$, this integral is evaluated as (33). Similarly for cases (ii) $\rho=1,\Delta=0$ and (iii) $\rho=1$, $\Delta>0$, it is evaluated as (34) and (35), respectively, since $\mathrm{e}^{-\sqrt{n}w}g_h(\mathrm{e}^{\sqrt{n}w})\le\mathrm{e}^{-\sqrt{n}w}$ holds for any w and

$$
\mathrm{e}^{-\sqrt{n}w}g_h(\mathrm{e}^{\sqrt{n}w})=\frac{h\eta(1+\mathrm{o}(1))}{\mathrm{e}^{h\eta}-1}
$$

holds for $w\le -n^{-1/4}$.

(See the next two pages for Eqs. (33)–(35).)

Now, combined with Lemma 7, it suffices to show that

$$
\begin{aligned}
\int_{-\infty}^{\infty}\mathrm{e}^{-\rho w}g_h(\mathrm{e}^w)\mathrm{d}w&=\int_0^\infty z^{-(1+\rho)}g_h(z)\mathrm{d}z\\
&=\frac{1}{\rho}\int_0^\infty z^{-\rho}\frac{\mathrm{d}g_h(z)}{\mathrm{d}z}\mathrm{d}z\\
&=\psi_{\rho,h}. \qquad (36)
\end{aligned}
$$

By letting $a=h\eta$ and $b=a/(\mathrm{e}^a-1)$, we can evaluate this integral as

$$
\begin{aligned}
&\int_0^\infty z^{-\rho}\frac{\mathrm{d}g_h(z)}{\mathrm{d}z}\mathrm{d}z\\
&=\int_0^\infty z^{-\rho-1}\frac{b\mathrm{e}^{-bz}-(a+b)\mathrm{e}^{-(a+b)z}}{a}\mathrm{d}z\\
&\quad+\int_0^\infty z^{-\rho-2}\frac{\mathrm{e}^{-bz}-\mathrm{e}^{-(a+b)z}}{a}\mathrm{d}z \qquad (37)
\end{aligned}
$$

(i) $\rho < 1$, $\Delta = 0$.

$$\iint \left(1-\frac{h(z_1,z_2)}{\sqrt{n}}\right)\frac{\mathrm{e}^{-(z_1,z_2)\Sigma_{01}^{-1}(z_1,z_2)^T/2}}{2\pi\sqrt{|\Sigma|}}\mathrm{e}^{-\sqrt{n}\rho z_1}g_h\left(\frac{\mathrm{e}^{\sqrt{n}z_1-z_2^2/2c_1}}{c_2\sqrt{n}}\right)\mathrm{d}z_1\mathrm{d}z_2$$

$$=\frac{(c_2\sqrt{n})^{-\rho}}{\sqrt{n}}\iint\left(1-\frac{h((w+z_2^2/2c_1+d_n)/\sqrt{n},z_2)}{\sqrt{n}}\right)\cdot$$

$$\frac{\mathrm{e}^{-((w+z_2^2/2c_1+d_n)/\sqrt{n},z_2)\Sigma_{01}^{-1}((w+z_2^2/2c_1+d_n)/\sqrt{n},z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}}\mathrm{e}^{-\rho w-\rho z_2^2/2c_1}g_h\left(\mathrm{e}^w\right)\mathrm{d}w\mathrm{d}z_2$$

$$\left(\text{by letting } \mathrm{e}^w=\frac{\mathrm{e}^{\sqrt{n}z_1-z_2^2/2c_1}}{c_2\sqrt{n}} \text{ and } d_n=\log c_2\sqrt{n}\right)$$

$$=\frac{(c_2\sqrt{n})^{-\rho}}{\sqrt{n}}\iint(1+\mathrm{o}(1))\frac{\mathrm{e}^{-(0,z_2)\Sigma_{01}^{-1}(0,z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}}\mathrm{e}^{-\rho w-\rho z_2^2/2c_1}g_h\left(\mathrm{e}^w\right)\mathrm{d}w\mathrm{d}z_2$$

$$+n^{-(1+\rho)/2}\iint_{\max\{|w|,|z_2|\}\ge n^{1/5}}\mathrm{e}^{-\rho w-\rho z_2^2/2c_1}g_h\left(\mathrm{e}^w\right)\mathrm{d}w\mathrm{d}z_2$$

$$\cdot\,\mathrm{O}\left(\sup_{w,z'}\left\{\left(1-\frac{h((w+(z_2')^2/2c_1+d_n)/\sqrt{n},z_2')}{\sqrt{n}}\right)\frac{\mathrm{e}^{-((w+(z_2')^2/2c_1+d_n)/\sqrt{n},z_2')\Sigma_{01}^{-1}((w+(z_2')^2/2c_1+d_n)/\sqrt{n},z_2')^T/2}}{2\pi\sqrt{|\Sigma_{01}|}}\right\}\right) \quad (32)$$

$$=\frac{(c_2\sqrt{n})^{-\rho}(1+\mathrm{o}(1))}{\sqrt{n}}\iint\frac{\mathrm{e}^{-(0,z_2)\Sigma_{01}^{-1}(0,z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}}\mathrm{e}^{-\rho w-\rho z_2^2/2c_1}g_h\left(\mathrm{e}^w\right)\mathrm{d}w\mathrm{d}z_2$$

$$+n^{-(1+\rho)/2}\iint_{\max\{|w|,|z_2|\}\ge n^{1/5}}\mathrm{e}^{-\rho w-\rho z_2^2/2c_1}g_h\left(\mathrm{e}^w\right)\mathrm{d}w\mathrm{d}z_2\cdot\mathrm{O}(1)$$

$$=\frac{(c_2\sqrt{n})^{-\rho}}{2\pi\sqrt{n|\Sigma_{01}|}}\int\mathrm{e}^{-z_2^2(\sigma_{00}/|\Sigma_{01}|+\rho/c_1)/2}\mathrm{d}z_2\int\mathrm{e}^{-\rho w}g_h\left(\mathrm{e}^w\right)\mathrm{d}w+\mathrm{O}\left(n^{-(1+\rho)/2}\iint_{|z_2|\ge n^{1/5}}\mathrm{e}^{-\rho w-\rho z_2^2/2c_1}g_h\left(\mathrm{e}^w\right)\mathrm{d}w\mathrm{d}z_2\right)$$

$$+\mathrm{O}\left(n^{-(1+\rho)/2}\iint_{|w|\ge n^{1/5}}\mathrm{e}^{-\rho w-\rho z_2^2/2c_1}g_h\left(\mathrm{e}^w\right)\mathrm{d}w\mathrm{d}z_2\right)$$

$$=\frac{(c_2\sqrt{n})^{-\rho}}{\sqrt{2\pi n(\sigma_{00}+\rho|\Sigma_{01}|/c_1)}}\int\mathrm{e}^{-\rho w}g_h\left(\mathrm{e}^w\right)\mathrm{d}w+\mathrm{o}(n^{-\frac{1+\rho}{2}})\,, \quad (33)$$

where (32) follows from

$$\left(1-\frac{h((w+(z_2)^2/2c_1+d_n)/\sqrt{n},z_2)}{\sqrt{n}}\right)\frac{\mathrm{e}^{-((w+(z_2)^2/2c_1+d_n)/\sqrt{n},z_2)\Sigma_{01}^{-1}((w+(z_2)^2/2c_1+d_n)/\sqrt{n},z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}}$$

$$=(1+\mathrm{o}(1))\frac{\mathrm{e}^{-(0,z_2)\Sigma_{01}^{-1}(0,z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}}$$

for (w,z_2) such that $\max\{|w|,|z_2|\}\le n^{1/5}$.

Here the first term is evaluated by integration by parts as

$$\int_0^\infty z^{-\rho-1}\frac{b\mathrm{e}^{-bz}-(a+b)\mathrm{e}^{-(a+b)z}}{a}\mathrm{d}z$$
$$=\frac{1}{\rho}\int_0^\infty z^{-\rho}\frac{(a+b)^2\mathrm{e}^{-(a+b)z}-b^2\mathrm{e}^{-bz}}{a}\mathrm{d}z$$
$$=\frac{\Gamma(1-\rho)}{\rho}\frac{(a+b)^{\rho+1}-b^{\rho+1}}{a}\,, \quad (38)$$

where we used the fact that for any $c>0$

$$\int_0^\infty\mathrm{e}^{-cz}z^{-\rho}\mathrm{d}z=\Gamma(1-\rho)c^{\rho-1}\,.$$

Similarly we have

$$\int_0^\infty z^{-\rho-2}\frac{\mathrm{e}^{-bz}-\mathrm{e}^{-(a+b)z}}{a}\mathrm{d}z$$
$$=\frac{1}{\rho+1}\int_0^\infty z^{-\rho-1}\frac{-b\mathrm{e}^{-bz}+(a+b)\mathrm{e}^{-(a+b)z}}{a}\mathrm{d}z$$
$$=\frac{1}{\rho(\rho+1)}\int_0^\infty z^{-\rho}\frac{b^2\mathrm{e}^{-bz}-(a+b)^2\mathrm{e}^{-(a+b)z}}{a}\mathrm{d}z$$
$$=\frac{\Gamma(1-\rho)}{\rho(\rho+1)}\frac{b^{\rho-1}-(a+b)^{\rho-1}}{a}\,. \quad (39)$$

Combining (37) with (38) and (39) we obtain (36) by

$$\int_0^\infty z^{-\rho}\frac{\mathrm{d}g_h(z)}{\mathrm{d}z}\mathrm{d}z$$

(ii) $\rho = 1$, $\Delta = 0$.

$$\begin{aligned}
&\iint \left(1 - \frac{h(z_1,z_2)}{\sqrt{n}}\right) \frac{\mathrm{e}^{-(z_1,z_2)\Sigma_{01}^{-1}(z_1,z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}} \mathrm{e}^{-\sqrt{n}\rho z_1} g_h\left(\frac{\mathrm{e}^{\sqrt{n}z_1 - z^2/2c_1}}{c_2\sqrt{n}}\right) \mathrm{d}z_1\mathrm{d}z_2 \\
&= (c_2\sqrt{n})^{-1} \iint \left(1 - \frac{h(w + (z_2^2/2c_1 + d_n)/\sqrt{n}, z_2)}{\sqrt{n}}\right) \cdot \\
&\qquad \frac{\mathrm{e}^{-(w+(z_2^2/2c_1+d_n)/\sqrt{n},z_2)\Sigma_{01}^{-1}(w+(z_2^2/2c_1+d_n)/\sqrt{n},z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}} \mathrm{e}^{-z_2^2/2c_1}\mathrm{e}^{-\sqrt{n}w} g_h\left(\mathrm{e}^{\sqrt{n}w}\right)\mathrm{d}w\mathrm{d}z_2 \\
&\qquad\qquad \left(\text{by letting } \mathrm{e}^{\sqrt{n}w} = \frac{\mathrm{e}^{\sqrt{n}z_1 - z^2/2c_1}}{c_2\sqrt{n}}\right) \\
&= (c_2\sqrt{n})^{-1} \frac{h\eta}{\mathrm{e}^{h\eta}-1} \iint_{w \le -n^{-1/4}} (1 + \mathrm{o}(1)) \frac{\mathrm{e}^{-(w,z_2)\Sigma_{01}^{-1}(w,z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}} \mathrm{e}^{-z_2^2/2c_1}\mathrm{d}w\mathrm{d}z_2 + \mathrm{o}(n^{-1/2}) \\
&= (c_2\sqrt{n})^{-1} \frac{h\eta}{\mathrm{e}^{h\eta}-1} \frac{1}{2\sqrt{|\Sigma_{01}|\left|\Sigma_{01}^{-1} + \begin{pmatrix} 0 & 0 \\ 0 & 1/c_1 \end{pmatrix}\right|}} + \mathrm{o}(n^{-1/2}) \\
&= (c_2\sqrt{n})^{-1} \frac{h\eta}{\mathrm{e}^{h\eta}-1} \frac{1}{2\sqrt{1 + \sigma_{11}/c_1}} + \mathrm{o}(n^{-1/2}) \,. \qquad (34)
\end{aligned}$$

(iii) $\rho = 1$, $\Delta > 0$.

$$\begin{aligned}
&\iint \frac{\mathrm{e}^{-(z_1,z_2)\Sigma_{01}^{-1}(z_1,z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}} \left(1 - \frac{h(z_1,z_2)}{\sqrt{n}}\right) \mathrm{e}^{-\sqrt{n}\rho(z_1 - \sqrt{n}\Delta)} g_h\left(\frac{\mathrm{e}^{\sqrt{n}(z_1 - \sqrt{n}\Delta) - z^2/2c_1)}}{c_1\sqrt{n}}\right) \mathrm{d}z_1\mathrm{d}z_2 \\
&= (c_2\sqrt{n})^{-1} \iint \left(1 - \frac{h(w + (z_2^2/2c_1 + d_n)/\sqrt{n}, z_2)}{\sqrt{n}}\right) \cdot \\
&\qquad \frac{\mathrm{e}^{-(w+(z_2^2/2c_1+d_n)/\sqrt{n},z_2)\Sigma_{01}^{-1}(w+(z_2^2/2c_1+d_n)/\sqrt{n},z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}} \mathrm{e}^{-z_2^2/2c_1}\mathrm{e}^{-\sqrt{n}(w-\sqrt{n}\Delta)} g_h\left(\mathrm{e}^{\sqrt{n}(w-\sqrt{n}\Delta)}\right)\mathrm{d}w\mathrm{d}z_2 \\
&\qquad\qquad \left(\text{by letting } \mathrm{e}^{\sqrt{n}w} = \frac{\mathrm{e}^{\sqrt{n}z_1 - z^2/2c_1}}{c_2\sqrt{n}}\right) \\
&= (c_2\sqrt{n})^{-1} \frac{h\eta}{\mathrm{e}^{h\eta}-1} \iint_{w \le \sqrt{n}\Delta - n^{-1/4}} (1 + \mathrm{o}(1)) \frac{\mathrm{e}^{-(w,z_2)\Sigma_{01}^{-1}(w,z_2)^T/2}}{2\pi\sqrt{|\Sigma_{01}|}} \mathrm{e}^{-z_2^2/2c_1}\mathrm{d}w\mathrm{d}z_2 + \mathrm{o}(n^{-1/2}) \\
&= (c_2\sqrt{n})^{-1} \frac{h\eta}{\mathrm{e}^{h\eta}-1} \frac{1}{\sqrt{1 + \sigma_{11}/c_1}} + \mathrm{o}(n^{-1/2}) \,. \qquad (35)
\end{aligned}$$

$$\begin{aligned}
&= \frac{\Gamma(1-\rho)}{\rho} \frac{(a+b)^{\rho+1} - b^{\rho+1}}{a} \left(1 - \frac{1}{1+\rho}\right) \\
&= \frac{\Gamma(1-\rho)}{1+\rho} \frac{\left(\frac{h\eta \mathrm{e}^{h\eta}}{\mathrm{e}^{h\eta}-1}\right)^{\rho+1} - \left(\frac{h\eta}{\mathrm{e}^{h\eta}-1}\right)^{\rho+1}}{h\eta} \\
&= \frac{\Gamma(1-\rho)}{h\eta(1+\rho)} \left(\frac{h\eta}{\mathrm{e}^{h\eta}-1}\right)^{\rho+1} \left(\mathrm{e}^{h\eta(1+\rho)} - 1\right) \\
&= \Gamma(1-\rho)\left(\frac{h\eta}{\mathrm{e}^{h\eta}-1}\right)^{\rho+1} \frac{\mathrm{e}^h - 1}{h} = \rho\psi_{\rho,h} \,,
\end{aligned}$$

where we used $\eta = 1/(1+\rho)$. □